Technical Report: NAVTRADEVCEN 4119-1

AD630260

THE USE OF CUING IN TRAINING TASKS: PHASE II

J. Annett
L. Paterson

University of Aberdeen
Scotland, United Kingdom
Contract N62558-4119

February 1966

DOD Distribution Statement

Distribution of this
document is unlimited.

THE USE OF CUING IN TRAINING TASKS: PHASE II

ABSTRACT

The report falls into three sections, a review of the literature on training for auditory tasks, an account of three experiments comparing cuing and knowledge of results as training techniques for a detection task, and the comparison of cuing and knowledge of results in an intensity discrimination task.

The review of the literature indicates some disagreement on the kind and amount of improvement in simple auditory tasks. Some improvement is undoubtedly due to familiarisation with the mechanics of listening and responding. Some may be due to changes in response criterion and some may be due to a genuine sensitisation to the auditory signals. The variety of techniques and performance measures does not facilitate straight-forward generalisation, although it is reasonably clear that some kindsof training can be effective.

Investigating a previous finding that cuing and knowledge of results affect response criterion differently, the subjects in a detection task were required to respond with three degrees of confidence. It was found that subjective confidence is not affected by training, but knowledge of results still produced more "risky" behaviour than cuing as defined by the distribution of detections and false positive responses.

Investigating the hypothesis that this difference was due to the necessarily higher rate of responding in knowledge of results an experimentally controlled rate of responding reduced the difference but did not eliminate it. A "cuing" procedure with no signal presented atall led to performance similar to that under "genuine" cuing, suggesting that in this task signal distribution is primarily what is learned and not, as had been hypothesized, the nature of the signal. Post-training vigilance performance appeared not to differ for the two techniques. In fact neither group showed significant changes in detection or false positives over a half hour vigilance session.

In the third section cuing, knowledge of results and reduced noise were compared in training intensity discrimination. In this case knowledge of results was effective, the other techniques not leading to improvement over five one-hour sessions. As in the detection task, however, cuing and knowledge of results were distinguished by increasing cautinn and increasing confidence respectively following training.

FOREWORD

Purpose

One of the critical needs of the Navy is to determine the "trainable factors" in sonar and to learn how to optimally conduct such training.

This study is the second phase of such an on-going program of experimentation. It is concerned primarily with the techniques of cuing and knowledge of results (KR) on a range of auditory detection and discrimination tasks.

In the first phase of this program (Annett and Clarkson, 1964), cuing was found to be more effective than KR in an auditory signal detection task. However, in view of the past history of learning theory it seems so unlikely that knowledge of results should turn out to be an inefficient training technique that its use was continued in the present phase of the study to be certain of the results.

Specifically, then, the main purposes of this phase were:

1. To investigate the surprising finding that cuing was found to be more effective than knowledge of results.

2. To clarify the finding that cuing and knowledge of results have different effects in auditory vigilance tasks, namely, that KR results in risky behavior (as shown by a large number of false detections), while cuing results in a more cautious approach (reduces false positive score).

Results

Among the results discussed in detail in the report are the following:

1. A critical review of the literature revealed extensive disagreement about the nature of auditory learning and about the optimal method of training.

2. The superiority of cuing over knowledge of results found in the previous phase for auditory detection was not confirmed in the present study, KR resulting in a greater, though insignificant, increase in detections.

3. The previous finding that cuing and KR result in different response effects was repeated. That is, KR led to a more lax response criterion, resulting in an increase in false positives, whereas cuing resulted in more cautious behavior, shown by a decrease in errors (false positives).

Implications

The results of this study are not yet applicable to sonar training. It is still necessary to extend this line of research to more complex signals, up to and including real sonar sounds. Also, longer periods than a half-hour are indicated for studying vigilance. However, this program has formed a baseline for the next phase of research using more complex discriminations.

The relative effectiveness of cuing versus KR as training methods may be a function of the specific task being learned. KR appears more advantageous when the task requires the detection of larger percentages of signals, regardless of the number of false responses. On the other hand, cuing gives better results when cautious behavior, that is, fewer false detections are desired.

There is also evidence that method of training interacts with stage of training. The determination of the optimal combination of training methods for type of task and stage of training requires further research. The next phase will investigate the effect of increased complexity on these interactions.

Gene S. Micheli

Gene S. Micheli, Ph.D.
Life and Behavioral Sciences
Department

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. LITERATURE REVIEW

1.1 Introduction.

In monitoring tasks such as sonar, man's role in the detection and classification of signals remains crucial. Despite highly developed hardware, certain perceptual functions are still performed better by man, especially when the input is complex or masked by noise. The sonarman's detection and classification skills are developed partly by training but also require two or three years of shipboard experience. Yet little is known about the learning process that takes place during either controlled training or experience on the task. The present project investigated mechanisms by which two training methods operate, as well as comparing their effectiveness. The methods used were cuing, which consisted of a prompt before each auditory signal to identify it or to warn the subject it was coming, and knowledge of results (KR) which was information on the correctness of a response given after the response had been made. Knowledge of results has theoretical support but cuing or guidance techniques do not feature in theoretical formulations of learning. Since cuing techniques are widely used they clearly need closer investigation.

It may be questioned whether improvements in perceptual judgments are really affected by practice and training. The sonar man appears to improve his perceptual skills by practice but we know little of the learning process involved. Before discussing training methods we must examine the evidence for saying that perceptual judgments, especially auditory, are susceptible to training.

1.2 Practice or Training Effects on Perceptual Judgments.

1.2.1. Absolute intensity thresholds.

Some doubt has recently been thrown on the effectiveness of practice in lowering the absolute intensity threshold. Swets and Sewall (1963) reached the "general conclusion" that the effect of practice is limited to the first session and that it is no more than 0 - 2 db for other than low frequencies. This conclusion, however, is limited to experiments in which feedback and signal specification were given. Signal specification is a means of ensuring that the subject knows the nature of the signal for which he is listening. Swets and Sewall specified the signal by giving 5

trials with the masking noise attenuated 10 db before each block of 100 trials. More complete specification has been used by Gundy (1961) who first presented the signal three times with its intensity raised 10 db above test level and then presented it at test level until the subject reported twice consecutively that he had heard it. Like Swets and Sewall, Gundy found that there was little practice effect when the signal was specified. Gundy specified the signal for one group but not for the other and then gave half of each group KR and the other hald practice. The non-specification group began at a lower level than the specification group and showed a gradual improvement while the specification group showed only a very slight improvement. The signal was then specified for all subjects and subsequently they all performed at a similar level. Some of the improvement found by other experimenters may, therefore, be due to subjects not knowing clearly what the signal was that they had to detect.

Other experiments quoted by Swets and Sewall do not entirely bear out the conclusion that there is only a small practice effect which is limited to the first session. Lukaszewski and Elliott (1961) using a 1000 c.p.s. tone compared the practice effect with and without KR. The mean threshold for the KR group was 3.4 db lower than the no-KR group. Most of this improvement occurred during the first session. For the group as a whole there was a 2 db improvement which extended over seven 35-minute sessions although most of it did occur in the first few minutes. Also with a 1000 c.p.s. tone Zwislocki et al. (1958) found a 2 db effect over four half-hour sessions.

The biggest improvements have been with low frequency signals. With a 100 c.p.s. tone, Zwislocki et al. (1958) found improvements of from 5 to 10 db with feedback or feedback plus bonus and 5 db without feedback or bonus. With feedback and bonus, a group given a forced-choice method combined with a tracking method, which involved changing the signal intensity to follow the subject's measured threshold, showed an improvement of 10 db. With the forced-choice procedure and feedback and bonus a 5 db effect was observed over eight blocks of 100 trials. The tracking procedure with feedback and bonus produced a 9 db effect over three half-hour sessions. The improvement is marked for all conditions and is greater with feedback than without. The length of time over which improvement continued varied considerably. For most conditions it continued over several sessions, but of the 10 db improvement for the combined forced-choice and tracking procedure 6 db was gained

in the first fifty trials. It may be that this method required a good deal of adjustment from the subject and that this large, rapid improvement was due to the adjustment taking place quickly.

These large improvements for low frequencies contrast with the rather small effects found for high frequencies. Zwislocki et al.(1955 and 1958) found a 6 db lowering of the threshold for a 125 c.p.s. signal but this was progressively reduced for tones up to 1000 c.p.s. Loeb and Dickson (1961) obtained results that are not in complete agreement with this, for although they did find that the effect at 1000 c.p.s. was smaller and less significant than the large effect at 125 c.p.s. they also found that at 500 c.p.s. the effect was insignificant. It was suggested that the large practice effect at low frequencies was due to the subjects learning to discriminate between the signal and a low frequency physiological noise. As predicted, no practice effect was found at any frequency when measured against a background of random noise. Efforts to induce a practice effect for a high-frequency tone by introducing a high-frequency background noise were unsuccessful. Moreover practice effects against a background of noise have been found. Annett and Clarkson (1964) found an increase in detection for both KR and cuing groups, and Gundy (1961) found an increase of 4 - 5 db with feedback but no signal specification and 1.5 db with feedback and specification. Both used a background of white noise. The relationship between noise background and the practice effect at different frequencies is not at all clear. So we cannot say why there are different degrees of improvement at different frequencies but there is an important distinction between the large improvements in thresholds with low frequencies and the rather small improvements at high frequencies.

Not all experiments on intensity thresholds show an improvement, however. Swets and Sewall (1963), as mentioned, found only a fairly small effect. There are several possible reasons for this failure to find an improvement. The effect of signal specification has already been discussed and it was suggested that some of the improvement which other experimenters have found may be due to the subject not knowing the signal characteristics clearly at the beginning. Since it is not always easy to specify the signal it is important to know that training with KR produces an improvement when the signal is unspecified.

Another possible reason for the lack of improvement found by Swets and Sewall is that in the first part of the experiment in which they used a two - alternative forced-choice procedure, they used four different signal energies in different blocks of trials. Campbell (1964), who also

found no systematic effect of practice with feedback, varied the loudness of the signal to keep average detection level at a predetermined value. It may be that when the amplitude of the signal is varied the subject has difficulty in establishing an appropriate criterion for reporting the presence of a signal. Some support for this suggestion is given by an experiment by Wiener (1964) who found in a visual task that training with greater or smaller signal amplitude produced change in performance during training but this did not transfer to a session with a middle value for amplitude. It would seem, therefore, that to increase subjects' ability to detect a particular signal of a particular amplitude, training should use that specific signal. This presents a difficulty if the problem is to train for detection of a variable signal.

In the second part of Swets and Sewall's experiment subjects had to give a "Yes/No" response on each trial and rate how confident they were about their response, after three days practice with this method. Again little improvement was found. It may be that some of the practice effect found by others was due to habituation to the experimental situation, whereas in the Swets and Sewall experiment this habituation has already taken place. Since in some cases the practice effect continued over several sessions it seems unlikely that this was the only improvement that took place.

1.2.2. Frequency.

Having discussed the research literature on intensity thresholds, we must now consider what evidence there is for improvement in thresholds for frequency. Unfortunately there is a lack of experiments on improvement of duration discriminations.

1.2.2.1. Frequency: absolute thresholds. Gibson (1953b) refers to experiments on the upper and lower limens for pitch. Guilford (1936) while measuring the lower limen by the method of minimal changes obtained data for which the difference between the first 10 and last 10 series was "almost significant enough to suggest a lowering of the limen as if by a practice effect during the course of the experiment". Humes (1930) found practice lowered the upper tonal limen and suggested that subjects become more discriminating in calling a sound "tone", as opposed to noise, after practice.

1.2.2.2. Frequency: differential thresholds. On frequency discrimination, a recent experiment by Campbell and Small (1963) found an improvement

over six one-hour sessions. For a KR and no-KR group combined there was a decrease from 7.3 c.p.s. to 4.2 c.p.s. in the median difference limen (DL). Also the constant error became less variable. Gibson's (1953b) review of perceptual learning mentions twelve studies on frequency discrimination of which ten show an improvement. Wyatt (1945) trained subjects for twelve 50-minute periods. They began by listening to a 500 c.p.s. tone and attempting to sing it. A stroboscopic technique gave the subject feedback. The subject then practiced on tones above and below the standard. Training was also given on direct frequency discrimination using a "progressive" technique with feedback. The progressive technique consisted of giving progressively more difficult discriminations. Wyatt thought that auditory imagery and motor participation were the most useful training aids but Gibson points out that these methods involved a great deal of feedback.

Other successful methods also used feedback. Seashore (1939) gave five to six hours progressive training with feedback with resulting lowered DL's and an increase of 38 points in centile rank on the Seashore test. Capurso (1934) found a gain of 24 centile points on this test following training with feedback and a technique of associating "mood" words with particular musical intervals. Gibson notes that the unsuccessful studies she mentions did not involve feedback.

Gibson asks whether training like that given by Wyatt results in reduced differential thresholds. Finer discrimination may be the result of learning specific names or categories for certain bands or ranges of stimuli. She suggests that perhaps differential reinforcement with pitch names reduces generalisation to the physiological limit.

1.2.3. Complex stimuli.

The previous discussion is concerned with improved judgments for simple stimuli but work has also been done on perceptual learning with regard to complex stimuli. Swets et al.(Swets, 1962; Swets, Harris, McElroy and Rudloe, 1964) used five - dimensional signals in experiments comparing various training methods. The amount of learning cannot be judged as pre-test levels are not given, but differences between groups on the post-test show that the treatments have had differential effects on performance. Sidley, Windgrad and Bedarf (1965) compared the effects of two kinds of KR on learning to identify complex sounds. With KR in which the signal remained audible during response and feedback, they found about 45% improvement in detections over ten hours, with no

sign of an asymptote being reached. With KR without signal-response overlap the improvement was about 20%.

Miller (1956) has compared the information transmitted in uni-dimensional and multi-dimensional judgments. For auditory, visual and tactile stimuli the mean was 2 - 6 bits for uni-dimensional discrimination and the average was rather higher for multi-dimensional. In a six - dimensional auditory experiment by Pollack and Ficks (1954) it was found that 7.2 bits were transmitted.

Gibson (1953a) reports some experiments on speech perception. Black (1946) found that when listeners were trained for hearing over an interphone system, greatest gains were produced by practice with the actual words used in test conditions; but Licklider and Pollack (1948) and Egan (1948) found an improvement in understanding masked or distorted speech over and above increased familiarity with the words. Gibson (1953b) also mentions data on intelligibility of spoken messages under conditions of constant noise level. When values for given words were corrected for word form, the correlation between frequency of correct transcription of sentences and average frequency of the rarest word was .77. Other training which gave some improvement was hearing the words with visual confirmation, and the same plus familiarity with the voice speaking. Important factors in speech perception in these experiments were learning the actual words used, average frequency of the words, visual confirmation and familiarity with the voice.

1.2.4. Summary

Although training methods used vary considerably and so does the amount of improvement and the number of sessions over which it extends it seems clear that improvements in auditory detection and discrimination have been found for intensity and frequency, and for complex stimuli. Although Swets' conclusion that improvement of intensity judgments is restricted to one session does not apply to all the experiments mentioned it seems quite likely that some of the improvement found is due to increased familiarity with the test situation. For example, Zwislocki et al. (1958) found a 6 db effect in the first 50 trials, and it seems possible that such rapid improvement could be due simply to increased familiarity with the task. Swets found only a small practice effect after a previous three days' practice with the rating method used, which would have made the subjects familiar with the task.

No experiment seems to have specifically attempted to measure the amount of improvement in auditory judgments which is due simply to increasing familiarity with the task, but there is one interesting, relevant experiment. Coover and Angell (1907) tested subjects for discrimination of shades of grey, then gave seventeen days of training on auditory discrimination learning and finally retested them on discrimination of the greys. They found that reaction times to the greys showed a transfer from the auditory discrimination. Although it is impossible to say how much of the improvement in the experiments mentioned is due to habituation to the experimental situation, when improvement continued for as much as seven 35-minute sessions (Lukaszewski and Elliott, 1961) or ten hours (Sidley, Windgrad and Bedarf, 1965) it seems as if it cannot all be attributed to increased familiarity with the task.

The Swets and Sewall (1963) and Gundy (1961) experiments using signal specification suggest that some improvement may be due to subjects not knowing clearly at the beginning of a session what the signal is that they have to detect. Since it may not always be easy to specify the signal exactly it is useful to know that performance improves in such a situation.

None of the experiments on complex stimuli necessarily show any increase in sensitivity to the stimuli but there are various ways in which training can improve performance by enabling subjects to learn about the task.

Having summarised evidence to show that improvements do take place in auditory judgments we must now look at the amount of improvement obtained with different kinds of training and how this improvement occurs. In particular, we must try to decide whether sensitivity to the signal is increased or whether the improvement is of some other kind.

1.3. Varieties of Training.

The previous section concluded that improvements in perceptual judgments may be obtained through practice or training. The next problem to be discussed therefore relates to differences in the amount and kind of change in performance following different training methods. The main training methods discussed are knowledge of results (the most-used) and cuing or prompting. Pretraining of various kinds, programmed instruction, and incentives have also been used.

1.3.1. Knowledge of results.

Knowledge of results has been defined as "knowledge which an individual or group receives relating to the outcome of a response or group of responses" (Annett, 1961). Annett discussed various techniques of KR which, although they all come within the definition of KR, differ widely in other ways. KR may, for example, be simply evaluative (e.g. saying 'Good') or an "end score", or some form of error feedback which might give the kind or amount of error. The response measures used vary as widely as the techniques of KR. The variety of forms of KR and response measures increases the difficulty of coming to general conclusions about KR and it is difficult to make a direct comparison of the results from different experiments.

Generally speaking, however, KR produces an improvement in performance. Some exceptions to this generalisation have already been discussed (Swets and Sewall, 1963; Campbell, 1964). The results of the relevant experiments are not all directly comparable since some experimenters have given results as threshold changes, others as latency, others as d', and others as percent detections, sometimes with percent false positives. Signal detectability theory (e.g. Swets, Tanner and Birdsall, 1961) has given increased importance to false positive rates. (Signal detection is treated as including a decision process by which the subject decides whether each subjective observation is to be judged a signal or no-signal.) The criterion for a "Yes" response may be lowered to increase detections but at the cost of an increase in false positives (FP's). But the rise in FP's for a particular rise in detections is less than if the subject were merely guessing. For those interested in monitoring tasks such as sonar the most appropriate measures are detection and false positive rates. Related to this is the question of whether any improvement following training is the result of a lowered response criterion or increased sensitivity. Before dealing with the results expressed as detections and FP's we will summarise the results of other types of experiments.

The threshold and latency experiments really only demonstrate that KR does improve performance on perceptual tasks. The experiments giving detection and FP rates are more interesting because of the insight they give into the mechanisms by which KR operates.

a. "Thresholds". These experiments have already been mentioned

in the discussion of intensity and frequency thresholds. Lukaszewski and Elliott (1961), Zwislocki, et al. (1958), Guilford (1936) and Campbell and Small (1963) have all found improvements, in some cases extending over several sessions. These experiments all dealt with intensity or pitch, however, but at least the important sound dimensions are trainable with KR. The reasons for Swets and Sewall's (1963) and Campbell's (1964) finding of little or no improvement have already been discussed.

b. Latency. This measure has been used by experimenters interested in response changes during a watch. Usually a decrement in latency occurs over time but KR of various kinds prevents or lessens the decrement during the sessions on which KR is given. Loeb and Schmidt (1963) gave both true and false KR on the latency of responses and found false KR lessened the decrement while KR eliminated it. Adams and Humes (1963) also found a lessened latency decrement with KR and this transferred to sessions with no KR. Since other groups were aware of the temporal distribution of the signals (detections were never below 98%) and neutral stimulation was ineffective, Adams and Humes interpret KR as "a habit operation for monitoring behavior." Their experiment rules out arousal or improved temporal expectancies as the method by which response latency was kept from increasing. The change in performance seems to be some kind of change in response criterion.

c. Detections and false positives. Although the above experiments illustrate that performance usually improves with KR, the most relevant results for our present project are those expressed in terms of detections and false positives since these are crucial to the problem of detection. If we accept the general finding of an improvement following KR then the next stage is to investigate the mechanisms by which KR operates. KR could affect sensitivity and response criterion independently. The experiments in Section 2 of this report investigate the hypothesis that KR lowers the response criterion. This hypothesis was based on the results of a previous experiment. It is convenient at this point to make a distinction between three kinds of KR:-

(1) full KR in which KR is given on all responses and also, if the situation is free-responding not fixed-response-interval, on missed signals.

(2) KR on response only, when the respond method is free-responding and signals may be missed. This is referred to here as partial KR.

(3) missed-signal KR in which information is given only about signals which are undetected. (This is not included in the original definition of KR but has been called KR by its users.)

In (2), partial KR, an increase in response-rate gives the subject more information about signal occurrence while in (1) and (3) it does not. In (1) if the response-method is fixed-interval the response rate cannot be raised and anyway full KR is already given; while with a free-response method full KR including missed signals is given regardless of response rate. In (3) the response rate can be changed but this does not alter the information given the subject since no feedback is given on responses made.

Using partial KR, Annett and Clarkson (1964) found an improvement in percent detections but also an increase in percent error. The hypothesis was made that this kind of KR encouraged subjects to respond more frequently to get information and that this higher rate of response carrie over into the post-test. One of the present experiments (Section 2) tests this by using fixed-response intervals so that the subject cannot increase his response rate and is already given full information. With this method there was less of an increase in detections and FPs than in the free-responding situation, as expected, but there was still some tendency for subjects to increase their total "yes" responses thus increasing detections at the cost of a slight increase in percent error. So it seems that KR may have a slight tendency to lower the response criterion apart from any necessity to do so to get information on signal occurrence.

This suggestion is supported by the results of the intensity discrimination experiment in the present report (Section 3). Here subjects had to classify comparison stimuli as "more," "less" or "same" compared with a standard. The group which had received KR increased the proportion of responses in the "more" and "less" categories while the cuing group increased responses in the "same" category. This was interpreted as an increased willingness to make an extreme judgment rather than to place unsure responses in the middle "same" category. It is as if subjects become more willing to test out their unsure judgments while receiving KR, and this lowered criterion for "yes" or extreme category responses carries over to the post-test. Thus partial KR has the effect of lowering the subject's response criterion so that he obtains more information, but even if full information is given the response criterion may be lowered, though perhaps to a lesser extent.

Relevant to the hypothesis of a lowered response criterion following partial KR is an experiment by Wiener (1963) who compared full KR, partial KR and a control condition. The two KR groups were about equal on detections but the full KR group had fewer FPs than the controls whereas the partial KR group had more. So full KR does not necessarily increase the FP rate even when detections increase, thus indicating that not all of the improvement following KR can be attributed to a change of criterion. The partial KR group, however, has a higher FP rate than the controls, as expected. In this instance partial KR lowered the response criterion as Annett and Clarkson (1964) found, but full KR did not lower the criterion.

It remains to see how missed-signal KR affects performance. When only missed-signal information is given the subject will not receive any more information if he increases his response rate, but the KR he is receiving tells him that if he wishes to detect all the signals he must increase his response rate. This is in fact what subjects do, although they are not told whether their increased numbers of responses are correct or not. Mackworth (1964) found that both true and false KR on missed signals gave more FP's and more detections than a control condition. Kinchla and Atkinson (1964a) found similarly that false KR on missed signals increased the probability of both a hit and a miss. So as well as lowering the response criterion to obtain more information, a subject may lower his criterion if he is told that he is missing some signals. This could have been why there was an increase in FPs and detections in the full KR, fixed-response-interval experiment reported in Section 2 and mentioned above. A wrong "No" response indicates to the subject that he has missed a signal. So far it looks as if KR does produce an increase in detections but at the cost of an increase in FPs which indicates that the response criterion has been lowered. The amount that the criterion is lowered and the way in which this change is brought about varies with the way in which KR is given.

To summarise these results,

a. Partial KR in a free-responding situation lowers the criterion (Annett and Clarkson, 1964; Wiener, 1963; Kinchla and Atkinson, 1964a; present report, Section 2). By comparison with full KR, which lowers the criterion less or not at all, it seems this lowering of the criterion is at least partly in order to obtain more information about signal occurrence, such information being dependent on responses when KR is partial (Wiener, 1963; present report, Section 2).

b. Even with full KR the criterion may be lowered though to a lesser extent than with partial KR (present report, Section 2). Also subjects become more willing to use extreme response categories (present report, Section 3). With full KR subjects do not receive more information by lowering their response criterion. It is as if subjects became more willing to test out their unsure judgments.

c. KR only on missed-signals lowers the criterion although no more information is obtained. Presumably KR tells subjects their response rate is too low and they increase it regardless of cost in FP's.

d. With full KR subjects may obtain more detections than controls but without a correspondingly higher FP rate (Wiener, 1963). This suggests that not all the improvement found with KR is due to lowered response criterion. KR, however, has also been found to produce apparently rather different results. Atkinson et al. (1964) found that if the signal was more likely to occur in one of two forced-choice intervals than in the other then KR increased the number of responses to that interval. Kinchla and Atkinson (1964b) found that subjects could utilise the sequential statistical properties of a signal schedule (in this case the probability of two consecutive trials with a signal presented in each) and that this effect was more marked with feedback than without. These results may be regarded as a more complex example of KR altering the response criterion, if the criterion is thought of as one that can be altered during the experimental session. Swets et al. (e.g. Swets, Tanner and Birdsall, 1961) tend to treat the criterion as fixed throughout a session, while Atkinson and Kinchla (1965) suggest there is a response bias which changes with learning during a session with feedback.

Generally speaking then, KR lowers the response criterion. One exception to this is a kind of KR not yet mentioned - KR on false positives only. Chinn and Alluisi (1964) used this type of KR and found a decrease in FP's and also an increase in reaction times for correct detections. Although an apparent contradiction to what has been said about the effect of KR, this makes sense when the actual information given is considered. The subject is told when he makes a wrong response and will therefore learn to decrease such responses without, however, learning about the true nature of the signal or of its temporal distribution. This tendency to decrease responses transfers to correct detections and shows as increased reaction times. The other results from Chinn and Alluisi's (1964) experiment unfortunately do not fit the above interpretation of the mechanism of KR. A group given KR on missed signals only gave a decreased FP rate with no significant change in detections - a completely

different result from other experiments with missed signal KR. Partial KR on correct detections only decreased the proportion of missed signals with no effect on FP's. This KR is informationally equivalent to partial KR on both correct and wrong responses which usually increases both detections and FPs. There is no obvious reason for these results differing from those previously described. Our description of how KR lowers the response criterion is not entirely satisfactory, therefore, and is in any case largely an ad hoc argument. Nevertheless as a working hypothesis it may be said that KR affects the response criterion, and that its effect on the criterion depends on what kind of information is provided by the KR. Usually the effect is to lower the response criterion.

1.3.1.1. KR and sensitivity. There seems no reason why KR should not affect sensitivity as well as the response criterion. Very few relevant experiments have used d', the signal detection theory measure of sensitivity (independent of response criterion).

The Swets and Sewall (1963) experiment in which it was claimed KR and practice had very little effect on d' has already been discussed and suggestions made as to why little effect was found. Swets and Sewall do admit that some of the increases in d' are not entirely negligible. Over five 2-hour sessions there was an improvement of .29 in d' (1.5 db). They say that of this about .5 to .75 db could be attributed to experimental error. Mackworth (1964) has also used the d' measure and found an increase for both false KR on missed signals and true KR on missed signals. There seems a definite although not very large increase following KR. One group was given a 40-minute session with no KR followed by two sessions with KR and a fourth without, one session per day being given. The increase in d' from the first to the second no-KR was .6 which is very approximately 3db. For a group given similar treatment with false KR instead of KR the increase was .3d' or approximately 1.5db. Mackworth suggests KR may enable subjects to learn the true nature of the signal which would certainly imply increased sensitivity. It is difficult to see, however, how false KR could affect sensitivity. Mackworth says that false KR may perhaps either increase arousal or help the subject to learn the probability of the signal, but the latter would imply a change of criterion not sensitivity while presumably physiological "arousal" could mean either. When the free-response method is used, d' cannot readily be calculated but for the fixed-response interval experiment reported here it has been calculated. An increase of .22 d' was found as a result of half an hour's training with KR. This is equal to approximately 1 db and is almost negligible, but it must be remembered that the training time was short. These results are far from conclusive but they do suggest that

KR may affect sensitivity directly. To the extent that this increase is due to learning about the characteristics of the signal itself, this improvement should transfer to a task where the signal rate and distribution differ from that during training - an important consideration in training for monitoring tasks.

1.3.1.2. How does KR affect sensitivity? The effect of KR on sensitivity has already been discussed. It is of interest both for understanding perceptual learning in general and for evaluating KR as a training method to try to decide how KR influences sensitivity.

Annett and Clarkson (1964) hypothesised that perceptual learning is "proportional to exposure to authenticated samples of the signal, and signal distribution information". This hypothesis was based on the results of a comparison of cuing and KR techniques, and other research literature on KR supports this hypothesis. Exposure to an "authenticated sample of the signal" does not simply consist of presenting the signal to the subject. It requires ensuring that he hears the signal and that it is identified or its occurrence confirmed.

The importance of KR giving a sample of signals which are authenticated was implied in an experiment by Sidley, Windgrad and Bedarf (1965) which also showed that identification which was continuous with the signal was more effective in enabling the subject to learn the nature of the signal and attach the correct response to it. In a complex discrimination task, Sidley et al. compared a form of KR in which the signal was continued during both response and feedback with a form in which response terminated the signal before KR was given. The KR with overlap between signal and KR was superior. They attribute this difference to a decay in the auditory memory during the delay before KR is given in the no-overlap condition.

Swets (1962) and Swets, Harris, McElroy and Rudloe (1964) found prompting techniques superior to overt responses followed by KR and concluded that streamlining of the task and temporal contiguity of sound and label are important factors in effective training methods. A follow-up experiment by Weisz and McElroy (1964) found that when KR was made simpler and given during presentation of the signal then KR was not inferior to prompting. This would prevent the subject forgetting what the signal had been before it was identified and help him to attach the correct label to it.

Gundy (1961) has shown the general importance of learning the characteristics of the signal in detection tasks. Gundy specified the signal for one group of subjects by presenting the signal at 10 db above test level and then at test level until the subject reported twice consecutively that he had heard the signal at test level. This ensured that the subject had heard the signal and that he knew it was the signal he had heard. This group and another for whom the signal was not specified were then given KR. The no-specification group began at a lower level than the other and then showed a gradual improvement while the specification group maintained a fairly stable level. The signal was then specified for all subjects and all then performed at about the same level. It appears that the no-specification group spent the KR session learning the characteristics of the signal and thus improving performance.

Thus we can conclude that learning the characteristics of the signal or signals is an important part of the learning in a perceptual task. KR can provide this necessary aspect of training, to varying degrees depending on the kind of KR, doing so most efficiently when there is temporal contiguity between the signal and its identification by KR.

If improvement in detection and discrimination tasks also involves learning the signal distribution, then full KR should lead to better discrimination than partial KR since full KR gives better information on signal distribution. Wiener (1963) found that full KR and partial KR give similar performance for percent detections but sensitivity or discrimination was better with full KR since there were fewer FP's than for partial KR. In the experiments reported here (Section 2) KR in the fixed-response interval condition (i.e. full KR) had a much better ratio of detections to false positives than KR in the free-response condition (i.e. partial KR).

Other evidence that learning the distribution of signals is an important factor comes from the Annett and Clarkson (1964) experiment. One method they used was "retrospective cuing" which consisted of a cue two seconds after each signal. The signals were near threshold and the problem was to detect them. A cue coming two seconds after the signal would not affect the number of signals heard although the cue was present for each signal. Thus the subject would not be given a larger sample of authenticated signals than would a partial KR group for which only responses are confirmed or disconfirmed. It would, however, give a better knowledge of the signal distribution. This "retrospective cuing"

group did in fact perform better than the partial cuing group.

Mackworth (1964) suggests KR enables subjects to learn the temporal distribution or signal characteristics, basing this suggestion on a comparison with false KR. False KR raises the over-all level of performance but KR also abolishes the decrement. Mackworth says the decrement is masked by learning but she cannot specify precisely what is learned.

Thus, as suggested by Annett and Clarkson, both opportunity to obtain an authenticated sample of the signal and to learn its distribution are probably important in perceptual learning and are provided by KR. The extent to which either is provided depends on the kind of KR. Insofar as the training effect of KR depends on learning about the signal itself and not merely the distribution this training effect should transfer to tasks with different signal distributions. Such transfer is essential since it is often impossible to train with the distribution that will occur in the task itself.

1.3.2. Cuing or prompting.

Cuing has been defined as "the provision of stimulus information before or during a response such that the response is made more effective or more likely to occur than would be the case without such information". Annett (1959) found that cuing proved to be an effective technique for both visual and auditory tasks. The superiority of cuing over KR was confirmed by Annett and Clarkson (1964) for an auditory task involving detection of a near-threshold signal. This experiment also showed completely different patterns of response following the two kinds of training - both showed an increase in detections but whereas KR showed an increase in percent error (FP's as percent of total responses), cuing showed a decrease. The present experiments which investigate further the difference between KR and cuing also find a similar pattern, although KR shows a greater increase in percent detections.

Swets (1961) compared a prompting technique with three kinds of feedback and found the prompting superior. The signals used were complex -- they could have either of two values for each of five dimensions. The task here was discrimination, not detection. With the prompting procedure, a computer typed out the values assumed by each of the five dimensions for a sound and then played that sound. The subject merely paced the presentation of each signal. The various kinds of feedback

required the subject to type the five numbers he believed identified the sound before being given feedback. A further experiment (Swets, Harris, McElroy and Rudloe, 1964) allowed subjects to choose between training methods for each trial, and the results supported the conclusions of the first experiment. Success was positively correlated with time spent on simple response to sound-label pairs and was negatively correlated with the proportion of time spent in active responding, receiving feedback and making second tries. Weisz and McElroy (1964) investigated the extent to which these results held for a visual task involving discrimination among three values for each of four dimensions of variation of unfamiliar forms. KR was made simple and was given while the signal was still present. In this case the superiority of prompting over KR disappeared. The best procedure was one "in which immediate feedback, cuing, and a simplified display of feedback information were combined."

1.3.2.1. Cuing and sensitivity. The first question to ask about these cuing techniques is whether they are effective because they change sensitivity to the signals or the response criterion. The Annett and Clarkson experiment and the experiments in Sections 2 and 3 of this report find that cuing both increases detections and decreases FP's which means that cuing has improved the ability to discriminate between signal and non-signal. Also the final ratio of detections to FP's suggests better discrimination than for KR. In the experiment in which fixed response intervals are used and d' can be calculated, cuing increases this measure by 1.26 whereas the increase for KR is only 0.22. This cuing group increase is equivalent to approximately 6 db.

Cuing appears also to affect the criterion but in the opposite way from KR. Detections are increased with no increase in FP's but instead a decrease. In the fixed-interval experiment (Section 2 of this report) beta can be calculated and was found to increase .22 for cuing (i.e. the criterion became stricter) while it decreased .14 for KR. In the intensity discrimination experiment (Section 3) KR showed an increased use of the extreme categories, but cuing increased use of the middle "same" category which can be thought of as a cautious response for unsure judgments. In this experiment, however, performance with cuing did not show a general improvement but reasons for this will be discussed later when comparing the usefulness of KR and cuing in different kinds of situations.

1.3.2.2. How does cuing affect sensitivity? As for KR, it is useful to discuss whether the results for cuing support the hypothesis that

learning is "proportional to authenticated samples of the signal, and signal distribution information."

Annett and Clarkson (1964) suggest that cuing was more effective than KR because it provided the maximum exposure to the signal and gave the fullest information about the distribution of the signal in the watch period. The cue, acting as a warning signal, allowed the subject actually to hear more of the faint signals than were heard by the other group. Howarth and Treisman (1958) have shown that a warning signal does in fact lower the threshold for the following signal. The results of the Annett and Clarkson experiment are consistent with the hypothesis that learning is proportional to exposure to authenticated samples of the signal and to signal distribution information. Cuing produced better discrimination than "retrospective cuing" (described in Section 1.3.1.2.) which would provide as good information about signal distribution but a smaller sample of heard and authenticated signals. Retrospective cuing was better than partial KR which would provide a similar sample of authenticated signals but would give poorer information on the signal distribution than retrospective cuing. The differential effectiveness of these three training methods corresponds to the extent to which they give authenticated samples of the signal and information on the signal distribution.

The Swets et al. (1962; 1964) and Weisz and McElroy (1964) experiments have been mentioned in the discussion on KR as illustrating the necessity of having signal samples clearly authenticated. The prompting technique enables the subject to see or hear what signal is being given the label, and to form the association between signal and label. With delayed KR the subject will have begun to forget the nature of the signal before it is identified for him. When the signal is still present during KR this method is as good as cuing.

The Gundy (1961) experiment using signal specification shows directly the importance of giving the subject authenticated samples of the signal. Gundy's method ensures that the subject hears the signal and knows it is the signal he has heard. Gundy points out that signal detection theory assumes an observer who "knows" what the signal is and that this assumption is not always met in training experiments. Where the signal is simple the kind of specification used by Gundy should ensure that the signal is "known" but when the signal is complex or variable it may be necessary to direct a good deal of training towards enabling the subject to recognise the signal.

In the experiment reported in Section 2 a group was included which was not given any presentations of the signal during training but merely the cue indicating when a signal should have been due. This group improved but the improvement was not significant whereas the cuing group did show significant improvement. This indicates that hearing the signal itself contributed to learning, but the no-signal group was small and the fact that their improvement was not significant does not rule out the possibility of some improvement due to learning the signal distribution.

Evidence showing generally that learning the temporal distribution of signals is an important factor in improvement in perceptual tasks of this kind has already been given in the section on KR. The retrospective cuing which Annett and Clarkson compared with KR showed this particularly.

The Annett and Clarkson hypothesis seems confirmed and it also may be concluded that cuing provides the necessary information on the signal and its distribution. The improvement with cuing should show transfer to other situations with the same signal to the extent that cuing has enabled subjects to learn about the signal itself. It is not so certain what effect the learning of a particular signal distribution or rate would have on transfer. Colquhoun and Baddeley (1964) compared various signal-rates for practice and test and found that a high practice signal-rate and a low test signal-rate gave the most within-session decrement on the criterion test, but Wiener (1963) found that mean signal detection rate was better with a high practice signal rate. A task such as sonar involves an extreme change from practice to task but it is impossible to train with a realistic signal rate, since even if it were known it would be very low.

1.3.2.3. Comparison of KR and Cuing. It appears that improvement in auditory detection and discrimination tasks is partly due to changes in response criterion and partly to changes in the ability to discriminate between signals or between signal and no-signal. The improvement in discrimination or sensitivity seems to be a matter of learning signal characteristics through exposure to authenticated signal samples and also learning the temporal distribution of the signal. Both KR and cuing provide these, though to varying degrees for different kinds of KR.

There are, however, some important differences between KR and cuing. In a detection experiment with near-threshold signals the cue acts

as a warning signal and lowers the threshold thus ensuring that the subject actually hears more signals than in other conditions. In a situation where the subject may not even be sure what he is listening for, cuing ensures that he has actually heard samples of the signal and that they have been authenticated. Prompting has been found more effective than KR in paired-associate learning and perceptual learning might be regarded as an instance of paired-associate learning Gibson (1953b) suggested that the experiments on absolute pitch discrimination demonstrate the applicability of the same kind of learning theory as can be applied to paired associates. In the Swets et al. (1962; 1964) and Weisz and McElroy (1964) experiments names had to be attached to large numbers of stimuli which the subjects had never tried to label in this way before. In these experiments prompting was better than, or at least equal to, KR.

One disadvantage of cuing is that cue-dependency may develop and the subject will respond automatically, paying little attention to the task. It is not usually meaningful to compare performance during cuing with the post-test since cuing ensures near-perfect performance. But Weisz and McElroy found that with a KR method which involved an element of cuing there was a decrement from training session to post-training test. In this experiment subjects had to name the value of each of the four dimensions of the stimulus to be identified. As each dimension was identified feedback was given, so that feedback on the first-named dimensions might help in correctly identifying the later dimensions. Some schedule of training which eliminates this decrement which occurs with a direct transfer from cuing to test is needed.

Knowledge of results (except for missed-signal KR), while specifying the signal less well than cuing, requires that the subject test out his own judgment before being given information, and therefore seems more suited to consolidating performance after the subject has learned what kind of stimulus to listen for or what responses are to be attached to various stimuli. KR has the advantage of avoiding cue-dependency but the low response criterion it encourages may be a disadvantage for some tasks. When long training seems necessary KR is also likely to be less boring than cuing since it requires more active participation by the trainee. These differences between the ways in which KR and cuing operate mean that their relative effectiveness will depend partly on the exact nature of the task. In experiments on detection of faint signals, cuing seems superior in terms of d' if not always in terms of detections. In this task the subject requires to be informed exactly what he is trying to hear and when it is likely to occur. Cuing, since it specifies the signal better than KR, seems the better method for such tasks.

In the intensity discrimination experiment reported here (Section 3), cuing was much poorer than KR. In this case signals were easily heard, the labels to be attached were easy to learn and large amounts of cuing were probably boring. In the Swets et al. and Weisz and McElroy experiments on the other hand, discrimination had to be made among many stimuli and the responses which were to be allocated were such that the subjects probably did not use them easily at first. In these cuing was superior or equal to KR.

Weisz and McElroy found some suggestion in their data that prompting was the best technique for subjects with low pre-test scores while KR was better for those with high pre-test scores. The difference in pre-test performance indicates that the subjects were required to learn different things about the task, and the different training methods presumably suited their different requirements. It seems likely that the different training methods will also be differentially effective at different stages of training. Cuing would probably be most useful in the early stages when the subjects are unsure about the nature of the stimulus or about which response to attach to which stimulus. KR would be more suitable later when the subject does not require so much assistance. Slight support is given for this though in a different kind of task by experiments by Carr (1930) on maze-learning. He found that guidance, which is comparable to cuing, was most effective in small amounts early in training while trial and error was more effective later and in larger amounts.

It is interesting too that the most successful method used by Weisz and McElroy combined KR and cuing. To get the advantages of both methods and to counteract their disadvantages there is probably some optimum combination of the two for any task, although subject variables would also influence the effectiveness of any method. The investigation of training schedules with these two methods seems a promising line of research.

1.3.3 Pretraining.

Several experiments have included some kind of treatment before the test session, not necessarily regarding it as a method of training but rather to test some hypothesis, for example about the effect of expectancy on signal detection or the related question of the decrement that occurs during a watch. These experiments have been grouped together here in order to consider whether these various pre-test treatments offer any possibilities as training methods.

Three experiments have investigated the effect of signal rate during practice on subsequent test performance. Colquhoun and Baddeley (1964) and Wiener (1963) using visual tasks reached the conclusion that training with a higher signal rate produced a higher rate of detections in the post-test, but Floyd, Griggs and Baker (1961) in an auditory task found that training with the same signal rate as used in the test gave the highest detection rate. In the Colquhoun and Baddeley comparison of different combinations of practice and test signal rates it was also found, however, that the transition from high practice rate to low test rate produced the greatest within-session decrement.

The advantages of using a high signal rate in training are not, therefore, well established and involve the possibility of increasing the decrement over time in the test session.

Within this section on pretraining may be included an experiment by Wiener (1964) on training with various signal amplitudes. His task was visual, the signal being an unusually large deflection of a voltmeter needle. Training was given with the test amplitude or a larger or smaller amplitude. Differences were observed during training but did not transfer to the test. The test, however, was held seven days later and it is likely that any training effect of one session would not be retained for so long. It has already been suggested that training with a variable signal or a signal other than the test signal, may not be very successful.

Of these experiments not even those concerned with signal rate offer any possibilities for training methods since it is necessary for tasks such as sonar to use a practice signal rate different from that found on the job.

1.3.4. Programmed learning.

Swets (1962) applied what he described as "the procedures of automated instruction" to the task of learning to identify multi-dimensional, non-verbal sounds. The procedures they used were continual interrogation and overt response, presentation of successive items conditional upon previous performance, learner-controlled pacing of the lesson, and so forth. They found these methods no better than "conventional" training methods and said that certain of the central features of automated instruction hindered learning in their task.

The application of these procedures in the Swets experiment may be questioned on some points as to whether they represent what would generally be considered the procedures of automated or programmed learning. Their first experiment compared a passive, prompting technique and overt responses followed by differing probabilities of reinforcement. To their surprise the methods with "reinforcement" were less effective than the prompting methods. The overt response method, however, did not minimise mistakes to ensure reinforcement but instead reinforcement often took place on a second try. It is general in programmed instruction to try to minimise errors, while the necessity of reinforcement through feedback is not accepted by everyone. Nor is the necessity of an overt response unquestionably established.

Swet's second experiment used a progressive technique by which the subject moved from easy to difficult discriminations. But this produced no better learning than methods in which the pairs for discrimination were chosen at random. The third session given the "progressive" group presented the pairs randomly and should have helped the transition to the test but nevertheless the method did not seem to have an advantage over others. This kind of progressive technique was used in some of the early experiments on judgments of pitch (Seashore, 1939; Wyatt, 1945) but their methods also included a good deal of feedback and the contribution made by the progressive technique cannot be estimated. The method has been used successfully with animals but it may simply enable them to learn the relevant dimension for discrimination.

The third experiment in this series was intended to test out a method of training which it was thought would maximise improvement. The subjects were trained progressively in the first session beginning with discrimination on the easiest dimension and on the easiest pairs within dimensions. A prompt was given and then the subject made an overt response with this in view. Discrimination was made more difficult until the sixth and seventh sessions, the pairs were chosen randomly and no prompt was provided but feedback was given. This long, complex training was not very successful. In terms of information transmitted, performance was no better than that obtained by Pollack and Ficks with more conventional methods although the experiment is not completely comparable. The percentages of detections in which all five dimensions were correctly identified was very low, from 9% to 1% for different subjects. The total number of sounds used was very high, 3125, and it may be unfair to judge

the success of the method without a direct comparison with other methods.

Swets next tried to estimate how well subjects could learn to identify sounds by their arbitrary names when trained under the procedures of automated instruction. They were also interested in replicating some of the previous results and so again compared a covert response prompting technique with overt response followed by feedback. The most effective technique was the prompting and even it led to only approximately 35% correct detections. As far as the comparison between prompting and overt response is concerned it should be remembered that prompting itself is an important aspect of programmed learning and so Swets et al.'s experiments cannot be taken as an indication that programmed learning is useless in this field. His results generally, however, are not very encouraging from the point of view of training for identification of complex sounds. Although several hours training were given perhaps it is not surprising that correct identification rates were low since large numbers of stimuli had each to be given absolute identifications. Short periods of training may be useful for estimating the effectiveness of various training methods but in this case the training was required to accomplish rather a lot in rather a short time.

A follow-up experiment by Swets, Harris, McElroy and Rudloe (1964) allowing the subject control of the course of training produced no better performance than when the experimenter determined the course of the lesson. It is hardly surprising that subjects would not know how best to train themselves for complex auditory discriminations. To say that this constitutes a failure of programmed instruction is unjustified since the aim of programmed instruction is to guide the student through a carefully arranged, pre-determined course of instruction. The general conclusion that "streamlining of the task and temporal contiguity of sound and label" are the most important factors for successful learning, accords fairly well with programmed-instruction techniques.

1.3.5. Incentives.

The incentives used in detection or monitoring tasks are usually money or the presence of an observer before whom the subject may wish to show his best performance.

An effect on the auditory threshold has been found following the introduction of a bonus after several experimental sessions. Lukaszewski and Elliott (1961) found that after twelve sessions the introduction of a

bonus had a 2 db effect on the threshold in the next two sessions. Zwislocki et al. (1958) found no effect when a bonus was introduced on the seventh session but there was a 2 db effect on the eighth. A small effect of a .9 db was found by Swets et al. after a bonus was introduced on the seventeenth day of training.

The effectiveness of a monetary reward on judgments of weight was compared with KR by Larimer and White (1964). No difference was found between the two although both were better than no knowledge. Then a session was given with no KR or reward but in which judgments had to be made in the presence of an anchor. The KR group showed contrast but the group which had been given reward showed a resistance to change in their judgments.

In a visual monitoring task, Bergum and Lehr (1964) gave monetary rewards for correct detections and penalties for missed signals. The rewarded group performed better than the controls in the first third of the session but no better in the remainder of it. In a second session without rewards, the incentive group was similar to the controls for the first third and significantly lower in the last third. So the effects of monetary reward may be short-term and when withdrawn may actually be a detriment to performance. The use of monetary incentives in training is therefore of questionable value.

The presence of an observer may also act as an incentive to better performance in detection tasks. Hardesty, Trumbo and Bevan (1963) found that observer-presented KR produced superior performance to machine-presented KR but it did not matter whether or not the observer was present in the room. Bergum and Lehr (1963) found that the presence of an officer in the room had a highly significant facilitatory effect on detection performance. They suggest that the conditions represent an extreme point along a dimension of perceived threat to the monitor. They compare their results with those of two other experiments for the ratio of errors between control and experimental groups. In an experiment (Bergum and Lehr, 1962) where monitors worked in pairs the ratio was .689. With the experimenter present it was .410 (Fraser, 1953) and when a military observer was present it was .363 (Bergum and Lehr, 1963).

It may be concluded that the presence of an observer, particularly one with some authority, facilitates performance but this is the kind of result which might not be replicated in a long-term study.

1.3.6. Conclusion.

The literature shows considerable diversity in the methods used and in the measures of training effectiveness, and few really satisfactory general conclusions can be drawn. However the following tentative conclusions emerge.

1. Most investigators have found improvements in auditory tasks as a result of practice. These tasks include detection of simple signals, the discrimination of pitch differences and the recognition of complex sounds.

2. In many studies training effects occur earlier rather than later in training and task familiarisation cannot be ruled out as a probable contributor to these results. This factor cannot, however, account for all improvement.

3. In detection tasks the subject may not initially be fully cognisant of all the relevant characteristics of the signal and some of the early training effects could well be due to learning the nature of the signal to be detected. Attempts to specify the signal are effective.

4. Knowledge of results, in various forms, is the most popular training technique and most investigators have obtained training effects with KR.

5. The results of several investigations suggest, however, that KR primarily affects the subjects' response criterion, generally leading to more risky behaviour. Sensitivity, as measured by d', is improved but this effect is generally not large.

6. KR is generally more effective when it overlaps with the auditory signal.

7. It has not been demonstrated that incentives have more than a temporary effect on performance in auditory tasks hence this aspect of knowledge of results would appear to be less important than its informative function.

8. Cuing, like signal specification and KR overlapping with the signal, is effective but in contrast with KR leads to more cautious behaviour and to increases in sensitivity.

9. In recognition tasks prompting has been found superior to some forms of KR whilst in at least one case of discrimination, prompting was less effective.

10. There is yet barely enough evidence to specify in general the conditions under which one or other of these two techniques will be superior, but it is reasonable to hypothesize that cuing may be better early in training and for signals difficult to specify by other means.

2. THREE STUDIES IN DETECTION TRAINING

2.1. Introduction

The three studies to be reported concern the effects for training on an auditory vigilance detection task, and arise directly from the previous studies by Annett (1959) and Annett and Clarkson (1964). The task, detection of short bursts of tone presented at irregular intervals against a background of white noise, may be taken as a highly simplified version of sonar watchkeeping.

Using a version of this task, Annett (1959) found a cuing technique, presenting a visual warning signal just before each tone to give positive transfer to an uncued standard task. Cuing was found to be more effective than either immediate or summary knowledge of results (KR) and more effective than practice with the background noise reduced. In that study a correction for false positives was applied to the detection scores, but in the second study (Annett and Clarkson, 1964) it became clear that cuing and KR were having different effects on correct detections and false positives. It was found that KR increased detections rather more than cuing but at the cost of an increased false positive rate, whereas cuing increased detections but decreased false positives. Wiener (1962) found a similar phenomenon in a visual detection task. Results were converted to a measure of information transmitted, and as with the first study using a correction for false positives cuing was found to be more efficient.

However, it remains that these two training techniques affect performance in different ways, and further investigation should reveal something of the learning mechanism. In the free responding vigilance situation with KR the subject can only get information by responding and the more he responds the more information he gets. As suggested by Wiener, the high rate of responding may carry over to the post-test. In cuing, by contrast, subjects do not have to respond in order to get information concerning the presence or absence of a particular signal and, since every signal is cued they have complete information about the signal distribution parameters. This latter information is denied subjects who only get right/wrong information after each response.

There is also the question of <u>what</u> is learned. The cue lowers the threshold to the auditory signal (Howarth and Treisman, 1958) and in this condition subjects should get maximum experience of the signal as such. This method permits them to build up a memory of the signal characteristics in a more efficient way than the KR technique, but as has already been pointed out, cuing also provides a potentially better indication of the signal distribution parameters.

One part of the present study investigates the hypothesis that cuing is effective by exposing the subject to authenticated samples of the signal enabling him to build up some kind of template of the characteristics of the signal. If this were so, cuing should be resistant to changes in the signal distribution parameters. Subjects should have no trouble in transferring from, say, a high signal frequency

to a low signal frequency. However, we have chosen to investigate this hypothesis in another, simpler, way by exposing the subject to the warning signal cues without the auditory signal. The distribution is the same, but subjects cannot gain any experience of the signal.

The second part of the study investigates the apparent effect of KR in increasing the false positive rate. If it can be argued that the free responding situation requires subjects under KR to lower their criterion in order to gain information, then by fixing the number of responses made the difference between cuing and KR in false positive rate should disappear. KR and cuing are therefore compared in a fixed observation interval situation where all other aspects of the task are the same as in the free responding situation.

A third aspect of the present study also concerns the effect of KR and cuing on response criterion. The higher detection and false positive rate under KR training could mean that this is making a subject adopt a more risky strategy rather than simply improving sensitivity to the signal. Cuing, on the other hand, induces a more cautious approach to the task, whilst increasing sensitivity to the signal. In the present series of experiments subjects are provided with three response keys (four in the case of fixed observation intervals) on which to register signals detected with one of three degrees of confidence, "sure," "fairly sure," and "unsure." The fourth key in the case of the fixed intervals experiment is to register "no signal." This additional method of recording responses is intended to show any changes in subjects' level of confidence as a result of training. This method is also relevant to the fourth aspect of the present series of experiments.

Comparing scores for the first and second half of the test Annett and Clarkson (1964) found some evidence of a vigilance decrement which was fairly marked in the case of the KR training. Any possibility of differential effects of training methods on vigilance decrement is worth exploring, so in the present investigation a number of subjects are given further training under either cuing or KR and are then tested on a half-hour watch. Broadbent and Gregory (1963) using graded response categories, report that some observers became more cautious during a watch period and this increased caution could show up as a vigilance decrement. In the present experiment it is of interest to see how the more risky performance following KR training holds up under prolonged watchkeeping as compared with the rather more cautious approach following cuing.

2.2. Apparatus

For all experiments in this series subjects are seated in an Industrial Acoustics Co. model 401A acoustic cabinet. The signals to be detected are generated by an oscillator tuned to 1,000 c.p.s. gated electronically to give a 100 millisecond burst of tone. The signals are mixed into the subject's headset with white noise at 50 db.

The signal intensity is attenuated to a near threshold value and is maintained throughout the experiment at this level, monitored in a sensitive valve voltmeter.

The signal presentation and response scoring are controlled by a punched tape program. Signals are delivered at irregular interv at an average rate of 4 per minute, the shortest interval being three seconds and the longest sixty seconds, the function relating inter-signal interval duration to frequency being approximately a negative exponential.

Facing the subject in the listening booth is a board with 4 coloured lights and 4 response keys. The lights provide cues, knowledge of results and also mark observation and response intervals. Their functions will be explained in detail in the discussion of experimental conditions. The four response keys indicate (from left to right) "yes, sure," "yes, fairly sure," "yes, unsure," and "no." The response keys are connected to nine electro-magnetic counters and to a multiple pen recorder. One counter records signals presented and a pair of counters for each of the four response keys counts detections and false positive responses. A detection is defined as any response made within two seconds of the arrival of a signal, this period being automatically timed. The pen recorder provides a check on the accuracy of the counters and also shows that the two second interval is appropriate, there being no clustering of responses later than two seconds.

Three experiments were carried out with this equipment. The first two experiments consisted of a pre-test, a period of training under one of several conditions and a post-test. For the third experiment some of the subjects for experiment 1 were recalled for a further test, further training and a vigilance session. Since some of the original subjects for experiment 1 were unable to return for experiment 3, additional subjects were collected and their results added to those of experiment 1.

2.3. Experiment 1

2.3.1. Comparison of Cuing, Knowledge of Results and No Signal

The aims of the first experiment were:-

(1) an attempt to repeat the finding that cuing and knowledge of results both improve detection but that cuing reduces the false positive score while knowledge of results increases it.

(2) to observe the effects of cuing and knowledge of results on the subjective confidence attached to "yes" responses before and after training.

(3) to test the hypothesis that cuing permits the subject to learn more about the nature of the signal against the hypothesis that cuing teaches primarily signal distribution parameters.

Using the pre-test, training, post-test design three groups of subjects were each submitted to one of the three following training

regimes.

1. Cuing

Exactly half a second before each signal an amber warning light was flashed. The light was a 100 per cent reliable indicator of the signal and subjects were instructed that after each cue light there would be a signal and that no signal would occur without the cue.

2. Knowledge of Results

Whenever a subject responded the green light flashed immediately if there had been a signal within the previous two-seconds period, but the red light flashed if there had been no signal.

3. No Signal

Using exactly the same schedule as in condition 1. the cue light was flashed at irregular intervals but no signal was given. Subjects were informed that during training no signals would be given but that the amber light indicated times at which signals would have occurred. The red and green lights were inoperative.

2.3.2. Responses

For conditions 1. and 2. three response keys were used and subjects were instructed that key 1 indicated "yes, sure," key 2 "yes, fairly sure" and key 3 "yes, unsure." In condition 1, subjects were instructed to respond according to their subjective certainty of hearing the signal which they, of course, knew would follow the amber cue light. Although, in principle, they were free to respond at any time responses were only made after the cue light. In condition 2. subjects were free to respond at any time and the Knowledge of Results lights, green for right and red for wrong, were independent of which degree of confidence was being expressed, that is which of the three "yes" keys was used. In condition 3, "no signal" subjects were instructed to press the fourth or "no" key after each cue light, thus ensuring that they attended at least to the lights, although they heard nothing but white noise on the earphones.

2.3.3. Pre and Post-test Conditions

All three groups were given the usual schedule of signals in white noise but without the benefit of cues or knowledge of results indicators. All three groups responded at will using the three "yes" keys.

For all groups the experiment was conducted in ten consecutive five-minute periods, periods 1 and 2 under pre-test conditions and periods 3 to 8 on one of the three training conditions. Between these periods there was a one minute rest pause during which the door of the listening booth was opened and subjects removed their earphones and relaxed. Following period 8 there was a five minute rest during which

the subjects were allowed to come out of the booth. Periods 9 and 10 consisted of the post-test under the standard test conditions and with the usual one-minute break between periods 9 and 10.

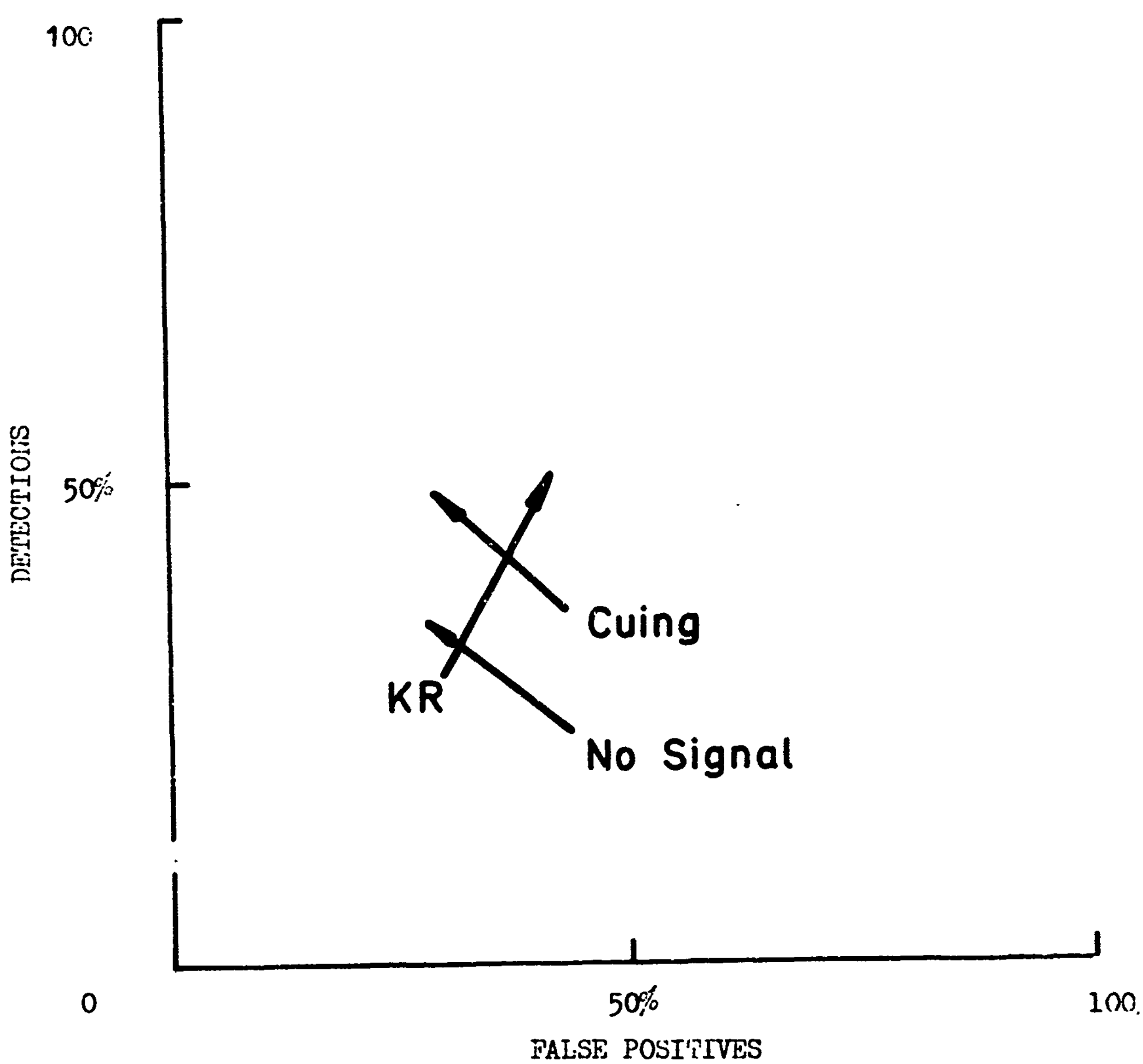
2.3.4. Experiment 1 - Results

With the three "yes" response categories we could look at the results in three ways, that is to say as if subjects were using a strict, medium or lax criterion. The preliminary analysis looks at the data from the point of view of the lax criterion that is to say all categories of "yes" responses are counted simply as "yes." The two main scores are thus % signals detected (% C) and % false positive response (% E), the latter being the ratio of wrong "yes" responses to total "yes" responses x 100. Table 1 summarises the data for the three training conditions on the pre-test and post-test for correct detection for %C (1a) and %E (1b). The data are also summarised in Figure 1. Although the pre-test scores appear to differ somewhat in both % C and % E the between groups differences are insignificant and the same is true of the post-test (see Tables 2a, 2b, 2c and 2d). Table 1 shows the predicted improvements in % C for Cuing and Knowledge of Results, the former being significant at $p < .0$ (one tailed) and the latter significant at $p < .005$ (one tailed). The apparent increase in the No Signal condition is not, however, significant but it should be remembered that this group has a rather small n so the possibility of some training effect without repeated presentation of the signal cannot be ruled out completely. Turning to the false positive responses, as predicted % E is reduced by cuing but $t = 1.68$ is insignificant on a one tailed test. Also as predicted Knowledge of Results increases false positive rate, $p < .025$ (one tailed). The No Signal group also shows a reduction in false positive rate but this is insignificant. In general the previous findings about Knowledge of Results and Cuing are confirmed although some of the predicted differences do not reach significance. The No Signal group reacts to training in a similar way to the Cuing group although the increase in detection rate and decrease in false positive rate are not significant. Bearing in mind that we are talking about a pattern even though the two elements of the pattern (detections up, false positives down) are not individually significant, it seems likely that this pattern is partly due to learning the distribution of signals. The advantage given by 40 minutes of actually hearing the signals in the case of the Cuing group is clearly not very great, and we must therefore attribute much of the learning which occurs in Cuing to knowledge of the signal distribution. This should not be too surprising since the signal itself is very simple, there is not much to learn about it and training effects are more likely to be found in the approach to the task.

Table 3 shows the distribution of responses made in the three response categories, for all yes responses in the pre-test and post-test. There are no very clear trends. We may note that the average number of responses mad decreases for Cuing, increases for KR and remains the same for No Signal. Both Cuing and KR make less use of the middle category and increase the proportion of "unsure" responses following

Figure 1

The effect of three training conditions on detections and false positives. The arrows originate at the pre-training score and terminate at the post-training score the barb indicating the direction of change.



training the pattern being very similar in both cases. Although it is possible to re-analyse the data using the medium or strict criteria a number of subjects would score 0 % correct in these categories. While this allows for an increase it does not permit one to calculate a satisfactory false positive rate; thus considering the loss in number of subjects which would be involved further analysis would appear to be not very fruitful.

2.4. Experiment 2

2.4.1. Free Response versus Fixed Observation and Response Intervals.

The aim of the second experiment was to test the hypothesis that the different response patterns found with Cuing and Knowledge of Results are due to the free response condition in which subjects under Knowledge of Results adopt a more risky criterion in order to learn about the signal and its distribution. It was argued that if subjects under Cuing and Knowledge of Results were required to make the same number of observations and the same number of responses differences in the effect of training on false positive rate should disappear. Again we have an opportunity to observe the effect of training on subjective confidence.

2.4.2. Conditions and Subjects

Four conditions were compared:-

1. Cuing with free response
2. Cuing with fixed observation and response intervals
3. Knowledge of Results with free response
4. Knowledge of Results with fixed observation and response intervals

The conditions for Cuing and Knowledge of Results were as in Experiment 1. However, under the fixed observation and response interval regime each five-minute period was divided into $7\frac{1}{2}$ second chunks as indicated by the white light on the subject's display. The white light remained on for 2 seconds and marked the interval during which subjects were permitted to indicate the presence or absence of a signal during the preceding $5\frac{1}{2}$ seconds when the white light was off. Subjects were required to respond on each occasion by pressing one of the three "yes" keys or the "no" key. The $5\frac{1}{2}$ second observation interval can be envisaged as divided into 11 $\frac{1}{2}$ second periods. A signal could occur at any of the half-second periods numbers 2 - 11. In the Cuing condition the amber cue light flashed $\frac{1}{2}$ second before each signal thus the first $\frac{1}{2}$ second of the observation interval had to be reserved for the cue when the signal was to occur in the first of its ten possible positions. In Cuing subjects were still required to respond during the two second response intervals as in Knowledge of Results. Since, in KR, feedback is contingent upon response, <u>KR was given only during the two second response interval</u>. For both free response and fixed interval conditions the actual schedule of signals used was the same. The experiment comprised a pre-test, training and post-test according to the same schedule as in Experiment 1.

For this experiment data from the free response Cuing and KR was taken from Experiment 1 and two further groups of ten subjects selected and paid in the same way comprised the fixed interval, Cuing and fixed interval, KR groups.

2.4.3. Experiment 2 - Results

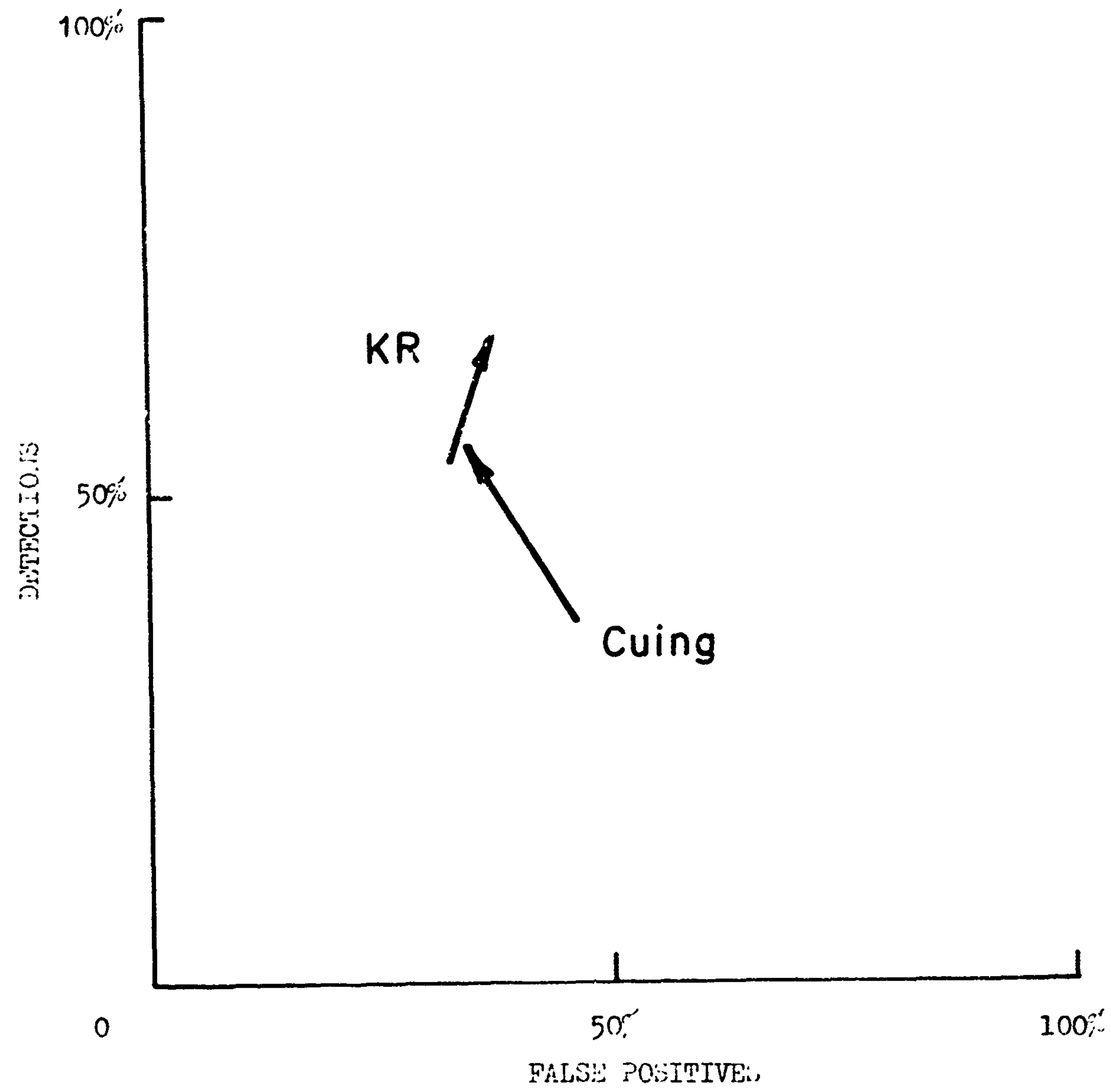
As for Experiment 1 the preliminary analysis counts all three response categories as "yes" responses. Table 4(a) shows % C for the four groups on pre and post-test and 4(b) % error for the four groups pre and post-test and Figure 2 summarises the results graphically. It can be seen from Table 4 and Figure 2 that even in the fixed interval conditions a similar pattern is found, with Cuing increasing detections and reducing false positives and KR increasing both detections and false positives. However the direction of the arrows on Figure 2 has been changed such that the tendency for KR to increase errors on post-test is somewhat reduced. Analysis of variance (Table 5) shows that the four groups do not differ significantly on either % C or % E on pre-test or post-test. However, looking at pre-test to post-test gains some tentative differences do emerge. Table 6 shows the analysis of variance for gains in % C and % E. There are significant differences between groups in % error which when further analysed are due to differences between Cuing and KR rather than between the two methods of responding, FR and FI. From individual 't' tests, Table 7 on the gains, it emerges that all four groups improved significantly on % correct, KR under the free response condition is the only group to show a significant increase in % E. Neither of the reductions in % E in the two Cuing groups is significant nor is the small increase in error in KR under fixed interval conditions. The result would, therefore, appear to confirm the prediction that when the rate of responding is controlled, as in the fixed interval condition, the difference between Cuing and KR in terms of false positives is reduced. The fact remains that the distinctive patterns produced by Cuing and KR are still present even in the fixed interval conditions, although to a less marked degree and it would be unwise to conclude from this evidence that differences between Cuing and KR can be entirely accounted for by the differential response rates acquired during training.

2.5. General Results from Experiments 1 and 2.

The previous finding that cuing and KR produce distinctly different patterns and may therefore be affecting learning through this mechanism clearly emerges from these results. We would like to describe performance in terms of a pattern of detections and false positives but statistical tests are only applied to each score, % C and % E individually. These considerations make it less easy to draw firm conclusions. We have found in Experiment 1 that giving the cue light without the signal produces a result characteristically like cuing, although slightly less efficient and not at all like KR. From this similarity, it would seem that whatever is learned under the Cuing condition it is not as had been hypothesized principally the

Figure 2

The effect of cuing and KR under fixed interval conditions.
The arrows originate at the pre-training score and terminate at the post-training score the barb indicating the direction of change



nature of the signal itself, but seems to be some appreciation of the signal distribution.

We have found that KR in the free responding situation increases the rate of responding (while in Cuing this is reduced even though performance becomes more efficient) and that when rate of responding is controlled the pattern is changed towards that of Cuing, but not entirely, so even in the fixed interval situation false positives are increased though not significantly so. With a much larger sample this small difference could be significant. We can only conclude that rate of responding has something to do with the difference between Cuing and KR but that it does not account entirely for the difference.

In both these experiments we measured three degrees of confidence in making "yes" responses. It transpired that training affected the degree of confidence only slightly and, oddly enough, the changes which did occur were very similar for both Cuing and KR. This result is unexpected in so far as the distribution of detections and false positives suggest that KR leads to a characteristically more risky approach, and that Cuing induces greater caution. This difference did not show up at all in terms of the subjective confidence attached to "yes" responses.

2.6. Experiment 3

2.6.1. Introduction

In a previous study, Annett and Clarkson (1964), there was some evidence of a vigilance decrement which was fairly marked in the case of KR training but only slightly following Cuing training. Such a difference, if confirmed, could be of considerable practical importance in training monitors. Having reached the conclusion that the main difference between the two methods was to be found in subjects' approach to the task the possibility arose that vigilance decrement might also be related to this. The cautious but efficient behaviour induced by Cuing might be maintained but what might be described as the optimism of subjects trained by KR could be reduced after a period with no KR.

Broadbent and Gregory (1963) applying signal detection theory to the analysis of vigilance, concluded that cautious performance becomes more cautious during a prolonged watch period while risky performance does not. This does not square well with Annett and Clarkson's results since increased caution should result in a reduction in detection rate. They found in fact that the higher detection rate (risky performance after KR) showed the greater decrement. However this finding is based on a comparison of the two halves of the ten minute post-test, thus confirmation is needed before either of the above mentioned results can be taken seriously.

2.6.2. Subjects and Conditions

In the present experiment subjects who had been trained previously under free response conditions with either Cuing or KR were invited to

Figure 3

% Correct detections for Cuing and KR as a function of time on watch

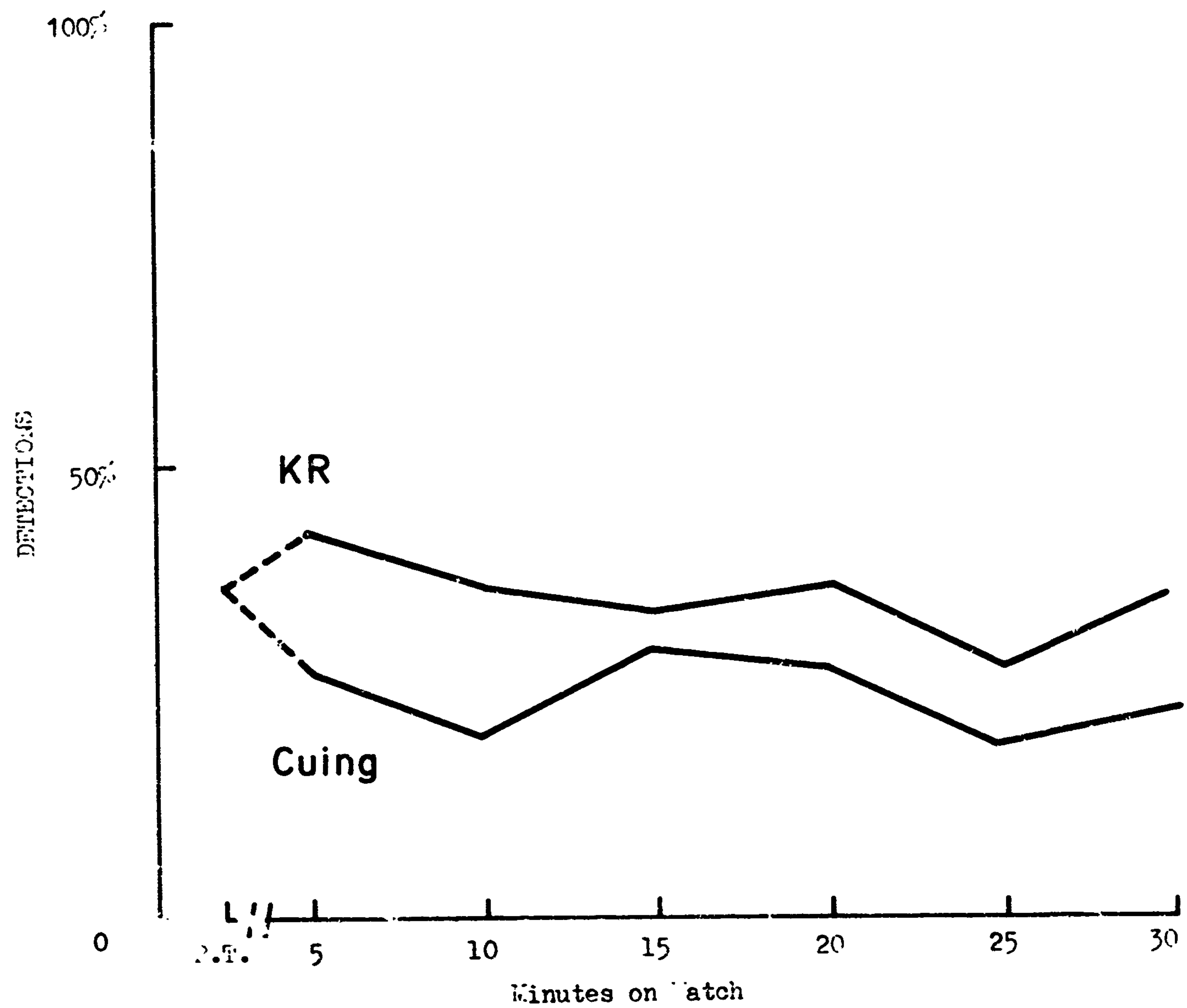
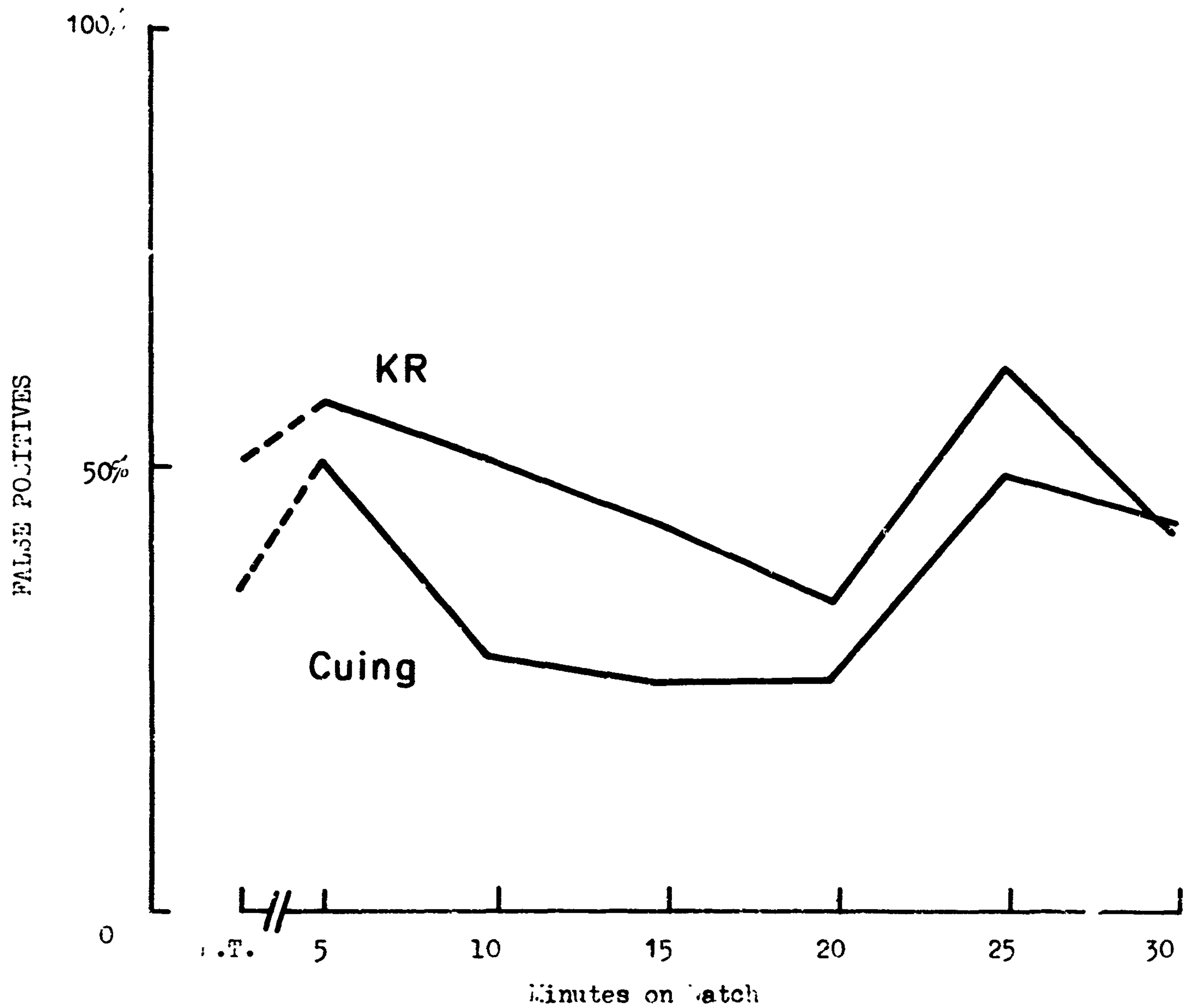


Figure 4

% E (False positives) for Cuing and KR as a function of time on watch



From this experiment there is no evidence of a differential effect of the two training methods on vigilance decrement. Both groups make slightly fewer responses at the end of the watch period and both groups become slightly less confident. Neither detections nor false positives show any significant decrement trend over six consecutive 5 minute periods and the difference between Cuing and KR in distribution of responses between correct detections and false positives seems to be preserved throughout the half hour watch period.

2. 7. Summary and Conclusions

2.7.1. Cuing and KR

It is now clearly established that in this task cuing and KR lead to characteristically different styles of performance. Both methods improve detection rates but while cuing also reduces false positive responses KR significantly increases them. Thus we can say that they have opposite effects on response criterion, cuing leading to cautious and KR leading to more risky behaviour. Since improvement in detection rates are comparable cuing can be regarded as more efficient except, of course, where the highest possible detection rate is required regardless of false positives.

2.7.2. What is Learned ?

By depriving subjects of signals but showing them when signals would occur we do not get significant increases in detections or reductions in false positives, but we cannot definitely assert that training is not possible under these conditions. The results fall into the same pattern as with the signal present and the differences in gains are not substantial. It seems probable therefore that under cuing what is learned about the signal distribution is at least as important as what is learned about the signal itself. This tentative conclusion is compatible with the general trend of these results that training is affecting functions such as expectancy and response criterion just as much, if not more, than it is affecting sensitivity to the signal as such.

2.7.3. Confidence Ratings

The results, as far as they have been analysed, have thrown up surprisingly little on the confidence rating of responses. Although we can characterize behaviour after cuing as cautious and after KR as risky both sets of subjects use the three degrees of confidence in about the same way when they are making "yes" responses.

2.7.4. The Effect of Controlled Rate of Responding

The hypothesis that KR leads to risky behaviour is partially confirmed by the result showing that when rate of responding is controlled KR does not lead to an increased false positive rate. However, the tendency does not appear to be entirely eliminated and it would be unwise to attribute the whole difference between cuing and KR to this source. The remaining difference could be due to

difference in the way the signal is perceived, Cuing producing a better opportunity to hear the signal than KR. It should be noted that under the fixed interval condition an average delay of 2½ seconds in the reception of KR does produce inferior learning, but not dramatically so.

2.7.5. Vigilance Decrement

In the vigilance experiment we have the somewhat unusual finding of no decrement over a half hour watch period. Just why this should be is unclear. It might be due to the preceding 25 minutes of testing and training. Yet a 5 minute rest period, usually found long enough to eliminate decrement, was put in with precisely this contingency in mind. A further possibility is the relatively low level of detections (less than 50% on average) even after training.

2.7.6. Differential Effects on Vigilance

The finding of no decrement in either group does not permit us to draw conclusions about differences between cuing and KR under conditions where decrement might occur. Thus although neither group showed a decrement, under other conditions, for instance a higher initial detection rate or an even longer watch period a differential decrement might be found. However, this possibility is somewhat reduced by the striking similarity between the two conditions on all counts. The curves for detection and false positives run virtually parallel through the 30 minutes and the slight evidence for changes in confidence are precisely the same for the two groups. Even the general differentiating pattern of detections and false positives for the two groups remains unchanged throughout the watch. Thus far then, we have failed to confirm the suggestion that KR might lead to a greater decrement. At the same time we have not found confirmation of an alternative hypothesis which could be argued from Broadbent and Gregory's results since there is no evidence that confidence levels change systematically with time on watch.

As a final comment it might be argued that although this task represents in a simplified form some aspect of sonar watch-keeping other aspects could be differently affected by the training variables chosen for study. Such generalisations as it is possible to make about cuing and KR for this task may not apply to others, e.g. the classification of complex signals.

3. A STUDY OF DISCRIMINATION TRAINING *

3.1. Introduction

The sonar man must learn to discriminate and identify a very wide range of complex sounds. On anecdotal evidence very high levels of skill can be attained but only after considerable practice. With material of such complexity and in the absence of any standard scale of performance one cannot be sure just what aspect of performance improves or to what extent, nor what kind of training would be most beneficial. Training could result in a wider repertoire of identifiable sounds, in finer discrimination between like sounds, greater use of circumstantial evidence in making decisions, in greater confidence, in greater proficiency in handling the equipment or in any combination of these and possibly other factors. In short, just what and how the sonar man learns is a complex problem.

One possible approach is to examine the kinds of discrimination which would appear to be necessary and, taking them at first individually, to test the effects of various kinds of practice and training. Gavin, Parker and Mackie (1959) in a report on "Trainable Factors in Sonar Operator Performance" have identified a number of "aural requirements" for various sonar operator functions. These include discriminations of relative loudness, duration and pitch in addition to more complex patterns. These authors stress that such discriminations are invariably masked by noise. The present investigation is the first of a series studying the trainability of discriminations relevant to sonar under noisy conditions. The experiment described here concerns the judgment of the relative intensity of sounds barely audible through masking noise.

Given the complex nature of the sonar man's task, this approach is open to the criticism that the learning of complex discriminations may not be deducible from the learning of simpler elements of the complex whole. This may very well be the case but the present experiments will provide a much needed base line from which to work on complex discrimination where the simpler elements are combined.

Indeed we begin these experiments with no great hope that the simpler discriminations are very susceptible to practice and the likelihood that sonar proficiency is due to higher level skills than these simple discriminations. The reader is referred to the literature review in Section 1 of this report.

The aim of this investigation is to find out if intensity discrimination is trainable. It should, however, be clear from the outset that a satisfactory answer can only be obtained if all reasonable training methods are employed and even then a possible successful training method

* This study was conducted with the assistance of J. Bloomfield, D. Marcer, G. Partington and V. Stanic; the help of Kingston High School in providing subjects is gratefully acknowledged.

could be overlooked and a negative answer cannot be considered as final, but only as an indication of the probability of finding a suitable method. This investigation goes further than most in comparing techniques with an unaided practice control.

KR in many training tasks is known to be profitable and yet in some studies it has turned out less effective than cuing, (c.f. Swets et al 1962, 1964, Annett and Clarkson 1964). These techniques are therefore included. In addition there is the possibility of varying signal/noise ratios for training purposes. Some sonar training is given on simulated sounds which are less noisy than the standard sea-going recordings. Cuing and noise reduction have in common the feature of reducing the difficulty of discrimination during training yet on the whole reduced noise tapes have been found less effective in transfer to the standard task, (Mackie and Harabedian 1964). Annett (1959) found comparable reduction in difficulty produced by cuing and reduced noise did not show comparable transfer to the standard task, cuing proving superior. In the first experiment these three conditions and the control are comparable in a simple pre-test/training/post-test design.

3.2. Apparatus and Conditions

Pairs of signals were presented against a background of continuous white noise at 50 db. Signal frequency remained constant at 1,000 c.p.s. and each signal lasted $\frac{1}{2}$ second the pair being separated by a $\frac{1}{2}$ second. S_1, the standard, was at 40 db. while S_2, the variable was of equal, greater or less intensity, three values of greater and three of lesser being used, $\pm$ .4 db, $\pm$ 1 db, $\pm$ 3 db.

Signals and noise were recorded on a Tandberg four-track recorder and played to subjects via padded Akai headsets, the output being monitored on a Marconi sensitive valve voltmeter.

Signal pairs were presented at regular 13 second intervals in sets of 48, a complete session consisting of four such sets separated by short rest intervals. "More," "Same" and "Less" signals were given in the proportions 3:2:3, the order being randomised over two adjacent sets or 96 pairs of signals.

Each pair of signals was introduced by a voice announcing "Item number - " followed $2\frac{1}{2}$ seconds later by S_1, in turn followed after $\frac{1}{2}$ second by S_2. There followed a 9 second gap for subjects to record their responses before the voice announced the next item.

Three tapes were made with white noise recorded on the second track. The output of both tracks was mixed into both ears through four padded headsets so that four subjects could be tested or trained simultaneously. Each subject had his own listening booth which, together with the padded headset, was adequate to exclude extraneous sounds. Subjects recorded their responses on numbered slips of paper as either "More," Less" or "Same."

The conditions so far described were used for pre and post-tests and as the training condition for a control group. Variations were

introduced into the tapes for the three further training conditions. Under Knowledge of Results a voice announced "More," "Less" or "Same" seven seconds after the termination of S_2 and two seconds before the announcement of the next item. Under the Cuing condition $2\frac{1}{2}$ seconds after the announcement of the item number a voice announced "More," "Same" or "Less" and S_1 followed two seconds later. 7 seconds elapsed following S_2 before the next item number was announced. Under the Reduced Noise training condition the masking noise was attenuated to a very low level but subjects practised as under test and control conditions without benefit of additional information.

For all groups the training schedule consisted of an hour-long session on each of five consecutive days, subjects attending at the same hour each day throughout the period. The signal pairs were presented in four batches of 48 with a short pause for rest between batches. On day 1 the first 96 pairs constituted the pre-test under the standard test condition. The second half of the session (96 pairs) was conducted under one of the four training conditions as were the whole of days 2, 3 and 4 and the first half of day 5. The final 96 pairs on Day 5 constituted the post-test.

Subjects were recruited from 7 - 18 year old pupils of both sexes at Kingston High School, Kingston upon Hull, in the free period following their summer examinations. All were briefly screened for hearing loss, using a Peters audiometer. Three showing losses of greater than 5 d.b. were rejected as a result of the test but two returned after having their ears syringed and passed the test. During the experiment three subjects failed to report for one or more of the practice sessions. The final groups comprised Control, N = 7, Cuing, N = 8 and Knowledge of Results, N = 6 and Reduced Noise, N = 7.

3.3. Results

The effectiveness of the four kinds of training can be simply described in terms of a single measure of efficiency, total number of correct responses out of 96 in the pre/test and post/test. On inspection Table 11 shows the biggest improvement due to KR, much smaller improvements for control and reduced noise and a small net loss for cuing. Analysis of variance, Table 12(a), shows pre-test differences between groups to be insignificant while on the post-test, Table 12(b) training groups differ significantly at the $p < .01$ level. Both pre-test and post-test were composed of two sets of 48 judgments. The first and second halves of the pre-test do not differ significantly but it was noted that most of the improvement in the KR group occurred in the second half of the post-test. Table 12(c) shows that the post-test differences are confined to the second half. Comparison of pre-test post-test gains reveals no significant differences except in the case of KR where the second half of the post-test is significantly better than the full pre-test score at the $p < .001$ level. Why the training effect should show only in the second half of the post-test is obscure, since the effect of KR is expected to be greatest immediately after training. There

is the possibility of a warm-up effect due to the transfer from training to testing conditions. All groups except control show lower post-test (first half) scores than were obtained either in the second half of the post-test or in the second half of the pre-test. However, since these differences are not significant no firm conclusion can be reached other than that KR is the only group showing a definite training effect.

In terms of a single index of efficiency only KR shows an improvement. The distribution of responses between the three response categories throws some light on the changes underlying this improvement. Table 13 shows the distribution of "Less," "Same" and "More" responses before and after practice, the actual distribution of signals being shown above in brackets. The main features of this table are that, as might be expected, most responses tend to be placed in the "Same" category but also that the "More" category is used least. After training the distribution for KR has moved in the direction of the actual distribution of signals, "Same" responses being reduced and "More" responses increased. On the other hand Cuing has moved even further from the correct distribution assigning fewer responses to both the "Less" and "More" categories and unlike any other group increasing the proportion of "Same" responses. The group data in table 13 is supported by the individual subjects' patterns of change which are shown in Table 16. The left hand column of table 16 shows the patterns of increase or decrease in the number of responses in the three categories from pre-test to post-test. The patterns at the top of the column show an increase in the number of responses in the middle category with a decrease in one or both of the extreme categories, while those at the bottom of the column show a decrease in the use of the middle category and an increase in one or both of the extremes. The most marked change towards the middle category (- + -) is at the top of the column and the most marked shift towards the extremes (+ - +) is at the bottom, other patterns being arranged in order of similarity to these. The patterns for individual subjects show that the Cuing group is clustered at the top half of the range, that is they tend to conform to the **pattern** (- + -). The KR group shows the opposite tendency, that is they are clustered at the bottom of the range where patterns indicate a shift towards greater use of extreme categories. The Reduced Noise and Control groups are scattered over the range so that they cannot be said to conform to any particular pattern.

These systematic changes in response distribution are consistent with a signal detection theory explanation in which Cuing tends to produce cautious behaviour and KR tends to produce more risky behaviour. Assuming that the "Same" category includes many of the doubtful responses and that after training these are re-allocated to the extreme categories, we should get a higher proportion of correct responses in the "Same" category. This appears to be happening in the case of KR as shown in tables 14 and 15. In table 14 there is a fairly big increase in the proportion of "Same" responses which are correct even though table 15

shows a smaller proportion of "Same" signals were correctly identified. At the same time the re-allocation of responses is not random since for KR higher proportions of "Less" and "More" responses were correct (table 14) and higher proportions of "Less" and "More" signals were clearly identified. In Cuing on the other hand although the increased number of "Same" responses results in a higher proportion of "Same" signals correctly identified the proportions for "Less" and "More" decrease (table 15).

The possibility that the KR group were simply matching their response distribution to the known signal distribution is worth considering since KR does provide this information. However, this is also true of the Cuing group which changes its response distribution in the wrong direction. Furthermore, as we have shown above the re-allocation of responses is not random. After training with KR subjects are more likely to allocate "More" signals to the "More" category and "Less" signals to the "Less" category, as would be expected on the signal detection theory of threshold. It seems unlikely therefore that subjects in the KR groups were reducing the number of "Same" responses simply to get a better match with the known frequencies of signals in the three categories, and one may reasonably conclude that a change in the level of acceptance of signals in the "Less" and "More" categories has occurred. Tables 13, 14, 15 and 16 can be interpreted as showing the following changes underlying the overall improvement in the KR group.

(1) All groups tend to under-estimate the frequency of signals in the "More" category.

(2) All groups tend to over-estimate the frequency of signals in the "Same" category.

(3) As a result of training KR moves towards the appropriate distribution of responses while Cuing moves in the opposite direction.

(4) These changes are consistent with a signal detection theory interpretation that under KR subjects become less cautious resulting in a higher proportion of "Less" and "More" signals correctly assigned. By contrast under Cuing subjects become more cautious and tend to use the "Same" category more frequently.

(5) The remaining two groups continue to under-estimate the number of signals in the "More" category and show no consistent change as a result of practice.

3. 4. Summary and Conclusions

This experiment has been successful in demonstrating that one of the discriminations believed to be relevant to sonar performance is trainable. With four hours of practice distributed over five days Knowledge of Results gives a highly significant improvement. In contrast to some other results Cuing seems to be ineffective in terms of gross efficiency. One might speculate that in this case continuous

cuing becomes boring and that smaller amounts of cuing perhaps interspersed with testing might be considerably more effective. Reducing the noise background seems to have very little effect by itself but again one might speculate that in conjunction with KR the technique could be of value.

A qualitative analysis of the response distribution before and after training indicates that while there are no obvious systematic effects of training in either Reduced Noise or Control groups, Cuing and KR exhibit opposite trends, the former becoming cautious and the latter becoming less cautious. This finding is of particular interest in view of the previous finding by Annett and Clarkson (1964) repeated by Annett and Paterson in Part 2 of the present Report. In a detection task Cuing also appeared to increase caution and KR to reduce it. This may, therefore, be a general feature of these two techniques, all the more interesting because theoretically they both provide the subjects with the same kind and amount of information about the task. In the detection task caution could have a certain merit depending on what weighting is chosen for false positives. In the method of scoring the present discrimination task, however, caution results in, if anything, a poorer post-training performance. The lack of improvement is however qualitatively different from that in the remaining two groups in so far as the data suggest a systematic change as a result of practice.

REFERENCES

1. Adams, J. A. and Humes, J. M. Monitoring of complex visual displays: IV. Training for vigilance. Human Factors, 1963, 5, 147-153.

2. Annett, J. The role of knowledge of results in learning: A survey. U. S. Naval Training Device Center Tech. Rep. No. 342-3, 1961.

3. Annett, J. Some aspects of the acquisition of simple sensori-motor skills. Unpublished D. Phil thesis, Oxford University, 1959.

4. Annett, J. and Clarkson, J. K. The use of cuing in training tasks. U. S. Naval Training Device Center Tech. Rep. No. 3143-1, 1964.

5. Atkinson, R. C., Carterette, E. C. and Kinchla, R. A. The effect of information feedback upon psychophysical judgments. Psychon. Sci., 1964, 1, 83-84.

6. Atkinson, R. C. and Kinchla, R. A. A learning model for forced-choice detection experiments. Psychology series, Institute for Mathematical Studies in the Social Sciences, Stanford University, Tech. Rep. No. 70, 1965.

7. Bergum, B. and Lehr, D. Vigilance performance as a function of paired monitoring. J. appl. Psychol., 1962, 46, 341-343.

8. Bergum, B., and Lehr, D. Effects of authoritarianism on vigilance performance. J. appl. Psychol., 1963, 47, 75-77.

9. Bergum, B. and Lehr, D. Monetary incentives and vigilance. J. exp. Psychol., 1964, 67, 197-198.

10. Black, J. W. Final report in summary of work on voice communication. U. S. Dept. Commerce Publ. Bd. No. 12051, 1946. (Office of Scientific Research and Development, 1945.)

11. Broadbent, D. E. and Gregory, M. Vigilance considered as a statistical decision. Brit. J. Psychol., 1963, 54, 309-323.

12. Campbell, R. A. Feedback and noise-signal detection at three performance levels. J. Acoust. Soc. Amer., 1964, 36, 434-438.

13. Campbell, R. A. and Small, A. M. Effect of practice and feedback on frequency discrimination. J. Acoust. Soc. Amer., 1963, 35, 1511-14

14. Capurso, A. A. The effect of an associative technique in teaching pitch and interval discrimination. J. appl. Psychol., 1934, 18, 811-818.

15. Carr, H. Teaching and learning. J. genet. Psychol., 1930, 31, 189-218.

16. Chinn, R. and Alluisi, E. A. The effect of three kinds of knowledge-of-results information on three measures of vigilance performance. Percept. mot. Skills, 1964, 18, 901-912.

17. Colquhoun, W. P. and Baddeley, A. D. Role of pre-test expectancy in vigilance decrement. J. exp. Psychol., 1964, 68, 156-160.

18. Coover, J. E. and Angell, F. General practice effect of special exercise. Amer. J. Psychol., 1907, 18, 328-340.

19. Egan, J. P. Articulation testing methods. Laryngoscope, 1948, 58, 955-991.

20. Floyd, A., Griggs, G. D., and Baker, R. Role of expectancy in auditory vigilance. Percept. mot. Skills, 1961, 13, 131-134.

21. Fraser, D. C. The relation of an environmental variable to performance in a prolonged visual task. Quart. J. Psychol., 1953, 5, 31-32.

22. Gavin, R. A., Parker, E. L., and Mackie, R. R. Trainable factors in sonar operator performance. U. S. Naval Training Device Center Tech. Rep. No. 306-1, 1959, (Confidential Report).

23. Gibson, E. G. Improvement in perceptual judgments as a function of controlled practice or training. Psychol. Bull., 1953, 50, 401-431. (a)

24. Gibson, E.J. A survey of research on improvement in perceptual judgments as a function of controlled practice and training. Human Resources Research Center, Air Research and Development Command, Lackland Air Force Base, Research Bull. 53-45, 1953 (b).

25. Guilford, J.P. Psychometric Methods. New York: McGraw-Hill, 1936.

26. Gundy, R.F. Auditory detection of an unspecified signal. J. Acoust. Soc. Amer., 1961, 33, 1008-1012.

27. Hardesty, D., Trumbo, D. and Bevan, W. Influence of knowledge-of-results on performance in a monitoring task. Percept. mot. Skills, 1963, 16, 629-634.

28. Howarth, C.I. and Treisman, M. The effect of warning interval on the electric phosphene and auditory thresholds. Quart. J. Exp. Psychol., 1958, X, 130-141.

29. Humes, J.F. The effect of practice upon the upper limen for for tonal discrimination. Amer. J. Psychol., 1930, 42, 1-16.

30. Kinchla, R.A. and Atkinson, R.C. The effect of false information feedback upon psychophysical judgments. Psychon. Sci., 1964, 1, 317-318 (a).

31. Kinchla, R.A., and Atkinson, R.C. The role of feedback in adaptive perceptual processes. NASA TMX-54, 059, Ames Research Center, Moffet Field, California, 1964 (b).

32. Larimer, G.S., and White, B.J. Some effects of monetary reward and knowledge of results on judgment. J. exp. Psychol., 1964, 67, 27-32.

33. Licklider, J.C.R., and Pollack, I. Effects of differentiation, integration and infinite peak-clipping upon the intelligibility of speech. J. Acoust. Soc. Amer., 1948, 20, 42-51.

34. Loeb, M., and Dickson, J. Factors influencing the practice effect for auditory thresholds. J. Acoust. Soc. Amer., 1961, 33, 917-921.

35. Loeb, M., and Schmidt, E.A. A comparison of the effects of different kinds of information in maintaining efficiency on an auditory monitoring task. Ergonomics, 1963, 6, 78-81.

36. Lukaszewski, J.S., and Elliott, D.N. Auditory threshold as a function of forced-choice techniques, feedback and motivation. J. Acoust. Soc. Amer., 1961, 33, 1046-1054.

37. Mackie, R.R., and Harabedian, A. A study of simulation requirements for sonar operator trainers. U. S. Naval Training Device Center Tech. Rep. No. 1320-1, 1964.

38. Mackworth, J. The effect of true and false knowledge of results on the detectability of signals in a vigilance task. Canad. J. Psychol., 1964, 18 (2), 106-117.

39. Miller, G.A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychol. Rev., 1956, 63, 81-97.

40. Pollack, I., and Ficks, L. Information of elementary multi-dimensional auditory displays. J. Acoust. Soc. Amer., 1954, 26, 155-158.

41. Seashore, R.H. Work methods: An often neglected factor underlying individual differences. Psychol. Rev., 1939, 46, 123-141.

42. Sidley, N.A., Windgrad, E. and Bedarf, E.W. Stimulus-identification overlap in learning to identify complex sounds. J. Acoust. Soc. Amer., 1965, 38, 11-13.

43. Swets, J.A. Learning to identify nonverbal sounds: An application of a computer as a teaching machine. U.S. Naval Training Device Center Tech. Rep. No. 789-1, 1962.

44. Swets, J.A., Harris, J.R., McElroy, L.S. and Rudloe, H. Further experiments on computer-aided learning of sound identification. U. S. Naval Training Device Center Tech. Rep. No. 789-2, 1964.

45. Swets, J.A. and Sewall, S.T. Invariance of signal detectability over stages of practice and levels of motivation. J. exp. Psychol., 1963, 66, 120-126.

46. Swets, J.A., Tanner, W.P. Jr., and Birdsall, T. G. Decision processes in perception. Psychol. Rev., 1961, 68, 301-340.

47. Weisz, A.Z., and McElroy, L.S. Response and feedback techniques for automated training of visual identification skills. U. S. Naval Training Device Center Tech. Rep. No. 789-3, 1964.

48. Wiener, E.L. Knowledge of results in a monitoring task. USAF AMRL Tech. docum. Rept. No. 62-82, 1962.

49. Wiener, E.L. Knowledge of results and signal rate in monitoring: A transfer of training approach. J. appl. Psychol., 1963, 47, 214-222.

50. Wiener, E.L. Transfer of training in monitoring: Signal amplitude. Percept. mot. Skills, 1964, 18, 104.

51. Wyatt, R.F. Improvability of pitch discrimination. Psychol. Monogr., 1945, 58, 2.

52. Zwislocki, J. Design and testing of earmuffs. J. Acoust. Soc. Amer., 1955, 27, 1154-1163.

53. Zwislocki, J., Maire, F., Feldman, A.S., and Rubin, H. On the effect of practice and motivation on the threshold of audibility. J. Acoust. Soc. Amer., 1958, 30, 254-262.

APPENDIX A

This appendix contains the sixteen tables referred to in the text.

Table 1(a)

Mean % Correct Detections for Cuing, KR and No Signal, pre-and post- tests

	C	KR	NS
N =	12	15	8
Pre-test	36.36	30.41	25.08
Post-test	49.27	50.50	36.00

Table 1(b)

Mean % Error for Cuing, KR and No Signal, pre-and post-tests

	C	KR	NS
N =	12	15	8
Pre-test	46.36	29.38	44.92
Post-test	50.23	40.96	23.36

Table 2 (a)

Analysis of Variance of Pre-test Scores, %C

Source	df.	S.o.S.	M.S.	F	p
Between groups	2	703.00	351.5	< 1	N.Sig.
Within groups	32	12625.96	394.56		
Total	34	13328.96			

Table 2 (b)

Analysis of Variance of Post-test Scores, % C

Source	df	S.o.S.	M.S.	F	p
Between groups	2	1212.46	606.23	2.21	N.Sig.
Within groups	32	8780.16	274.38		
Total	34	9992.62			

Table 2 (c)

Analysis of Variance of Pre-test Scores, % E

Source	df.	S.o.S.	M.S.	F	p
Between groups	2	1594.54	797.27	1.74	N.Sig
Within groups	32	14634.54	457.329		
Total	34	16229.08			

Table 2 (d)

Analysis of Variance of Post-test Scores, % E

Source	df.	S.o.S.	M.S.	F	p
Between groups	2	1146.09	573.05	1.5	N.Sig
Within groups	32	12027.74	375.87		
Total	34	13173.83			

Table 3

Distribution of "Yes" Responses on 3 Confidence Categories

pre and post-test

		"Sure"	"F.Sure"	"Unsure"	Average Number of Responses
CUING	Pre	43.44%	34.16%	22.39%	30.50
	Post	46.93%	21.16%	31.91%	27.17
KR	Pre	45.58%	29.50%	24.93%	20.34
	Post	42.47%	19.24%	38.29%	31.86
No Signal	Pre	34.05%	40.06%	25.89%	19.47
	Post	33.33%	41.67%	25.00%	18.00

Table 4(a)

% Correct Detections for Cuing and Knowledge of Results under Free Response and Fixed Interval Response Conditions Pre and Post-test

% C

		C	KR
F.I.	Pre	36.30	53.40
	Post	55.19	66.37
F.R.	Pre	36.86	30.41
	Post	49.27	50.50

Table 4 (b)

% E.

		C	KR
F.I.	Pre	44.5	32.14
	Post	34.41	36.47
F.R.	Pre	42.36	29.88
	Post	30.23	40.96

Table 5 (a)

Analysis of Variance of Cuing and KR under Free Response and Fixed Interval Conditions

% C Pre-test

Source	df.	S.o.S.	M.S.	F.	p
Between groups	3	3274.96	1091.65	1.863	N. Sig.
Within groups	42	24604.57	585.82		
Total	45	27879.54			

Table 5 (b)

% C Post-test

Source	d.f.	S.o.S.	M.S.	F.	p
Between groups	3	1981.09	660.36	1.559	N. Sig.
Within groups	42	17789.40	423.56		
Total	45	19770.49			

Table 5 (c)

% E Pre-test

Source	d.f.	S.o.S.	M.S.	F	p
Between groups	3	1836.28	612.09	1.76	N.Sig.
Within groups	42	14645.77	343.71		
Total	45	16432.05			

Table 5 (d)

% E Post-test

Source	d.f.	S.o.S.	M.S.	F	p
Between groups	3	793.30	264.4	< 1	M.Sig.
Within groups	42	16159.20	384.7		
Total	45	16952.50			

Table 6(a)

Analysis of Variance of Pre-test to Post-test Gains

%C

Source	d.f.	S.o.S.	M.S.	F	p
Between groups	3	569.53	189.84	< 1	N.Sig.
Within groups	42	20205.02	481.07		
Total	45	20774.55			

Table 6(b)

%E

Source	d.f.	S.o.S.	M.S.	F	p
Between Cuing and KR	1	4197.73	4197.73	9.745	< .01
Between FR and FI	1	148.80	148.80	< 1	
	1	131.54	131.54	< 1	
	42	18091.93	430.76		
Total	45	22570.00			

Table 7(a)

Pre-to Post-test Gains

%C	Gain	t	p
Cuing FR	12.40	2.102	< .05 (one tailed)
KR FR	20.90	3.24	< .005 (one tailed)
Cuing FI	18.88	2.21	< .05 (one tailed)
KR FI	12.97	2.59	< .05

Table 7(b)

% E	Gain	t	p
Cuing FR	- 12.13	1.68	N.S.
KR FR	+ 11.08	2.46	< .025
Cuing FI	- 10.13	1.23	N.S.
KR FI	+ 4.33	< .1	N.S.

Table 8(a)

Mean % C for Cuing and KR

	Pretest	1	2	3	4	5	6
Cuing	37.14	28.15	26.73	30.99	23.76	19.57	23.13
KR	37.33	44.43	37.82	34.35	37.71	28.54	36.58

Table 8(b)

Mean % E for Cuing and KR

	Pretest	1	2	3	4	5	6
Cuing	39.94	51.67	29.33	26.13	27.70	50.62	45.19
KR	51.03	58.24	52.13	44.41	35.03	62.51	43.65

Table 9(a)

Analysis of Variance by 5 Minute Periods

% C Cuing

Source	d.f.	S.o.S.	M.S.	F	p
Trials	5	1107.22	221.44	2.48	N.Sig.
Subjects	9	6622.94	735.88		
Interaction	45	4016.63	89.26		
Total	59	11746.79			

Table 9(b)

% C KR

Source	d.f.	S.o.S.	M.S.	F	p
Trials	5	1340.26	268.05	2.46	N.Sig.
Subjects	9	23050.11	2561.12		
Interaction	45	4898.30	108.85		
Total	59	29288.66			

Table 9(c)

% E Cuing

Analysis of Variance by 5 Minute Periods

Source	d.f.	S.o.S.	M.S.	F.	p
Trials	5	7189.07	1437.81	3.46	< .025
Subjects	9	5248.28	583.14		
Interaction	45	18715.72	415.90		
Total	59	31153.07			

Table 9(d)

% E KR

Source	d.f.	S.o.S.	M.S.	F.	p
Trials	5	5221.41	1044.28	2.38	N.S.
Subjects	9	9579.67	1064.41		
Interaction	45	19731.47	438.48		
Total	59	34532 .55			

Table 10

% Responses in Three Categories

Cuing	"Sure"	"Fairly sure"	"Unsure"	Average no. of responses
Pre-test	37.23	30.66	32.12	27.4
1st 10 minutes	47.2[illegible]	27.36	25.37	2[illegible].1
Last 10 minutes	32.00	35.33	32.67	15.0

KR	"Sure"	"Fairly sure"	"Unsure"	
Pre-test	39.31	18.21	42.48	34.6
1st 10 minutes	45.57	16.14	38.29	32.6
Last 10 minutes	37.89	15.25	46.86	25.6

Table 11

Mean Number of Correct Responses
on Pre-test and Post-test

	Control	Reduced Noise	K. of R.	Cuing	Max. Score
Pre-test	36.57	41.25	38.83	40.88	96
Post-test	40.14	43.00	49.5	39.25	96

Table 12

(a) Analysis of Variance of Pre-test Scores

Source	df.	S.o.S.	M.S.	F	p
Between Groups	3	60.70	20.23	1.39	N.S.
* Within Groups	52	758.14	14.58		
Total	55	818.84			

(b) Analysis of Variance of Post-test Scores

Source	df.	S.o.S.	M.S.	F	p
Between Groups	3	221.85	73.95	5.60	$< .01$
* Within Groups	52	686.14	13.195		
Total	55	907.99			

(c) Supplementary Analysis, 1st and 2nd halves of the Post-test

Source	df.	S.o.S.	M.S.	F.	p
Between Groups (1st half)	3	70.2	23.40	1.81	N.S.
Between Groups (2nd half)	3	168.2	56.07	4.34	$< .01$
Within Groups	96	1241.4	12.93		

* Note: Pre and Post-tests were both divided into two halves hence $N = 28 \times 2 = 56$

Table 13

% Responses in "Less", "Same" and "More Categories, Pre-and Post-test

		L	S	M
Distribution of Signals		(37.5)	(25.0)	(37.5)
Reduced Noise	Pre	32.69	44.79	23.53
	Post	34.90	42.58	22.53
Knowledge of Results	Pre	28.99	46.70	24.30
	Post	32.29	28.47	38.92
Cuing	Pre	38.94	35.03	25.92
	Post	31.51	44.01	24.48
Control	Pre	23.66	59.82	16.52
	Post	37.80	48.22	13.99

Table 14

% Correct Responses in "Less", "Same" and "More" Categories, Pre-and Post-test

		L	S	M
Reduced Noise	Pre	56.18	26.74	56.08
	Post	57.85	28.45	55.48
Knowledge of Results	Pre	53.90	25.65	52.85
	Post	59.13	37.21	55.75
Cuing	Pre	51.50	26.76	50.76
	Post	55.80	27.81	45.23
Control	Pre	57.86	26.12	53.15
	Post	55.11	26.85	57.41

Table 15

Signals Correctly Assigned to
"Less", "Same" and "More" Categories
Pre-and Post-test

		L	S	M
Reduced Noise	Pre	48.97	47.92	33.69
	Post	53.83	48.46	33.33
Knowledge of Results	Pre	41.67	47.92	34.25
	Post	50.92	42.38	58.33
Cuing	Pre	53.47	37.50	35.08
	Post	46.89	48.96	29.53
Control	Pre	36.50	62.50	23.42
	Post	55.56	51.79	21.42

Table 16

Patterns of Pre-test to Post-test

Change of Response Distribution

Pattern L S M	Cuing	KR	Reduced Noise	Control
- + -	*****		***	*
- + +	*			
+ + -			*	*
0 + -		*		
+ 0 -				*
- - +	**	***	*	
+ - -				*
+ - +		**	**	***

0 = No Change

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) University of Aberdeen Scotland, United Kingdom	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b GROUP

3. REPORT TITLE

THE USE OF CUING IN TRAINING TASKS: PHASE II

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report (Final) 30 June 1964 - 1 October 1965

5. AUTHOR(S) (Last name, first name, initial)

Annett, John
Paterson, Laura

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
February 1966	v., 74	53

8a. CONTRACT OR GRANT NO. N62558-4119 b. PROJECT NO. 7570-3 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report NAVTRADEVCEN 4119-1 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES None	12. SPONSORING MILITARY ACTIVITY Life and Behavioral Sciences Department U.S. Naval Training Device Center Port Washington, New York 11050

13. ABSTRACT

This report falls into three sections: a review of the literature on training for auditory tasks, an account of three experiments comparing cuing and knowledge of results as training techniques for a detection task, and the comparison of cuing and knowledge of results in an intensity discrimination task.

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Cuing Feedback Knowledge of Results Training Auditory Detection Auditory Discrimination						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ______________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ______________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through ______________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.

(back)

Unclassified

Security Classification